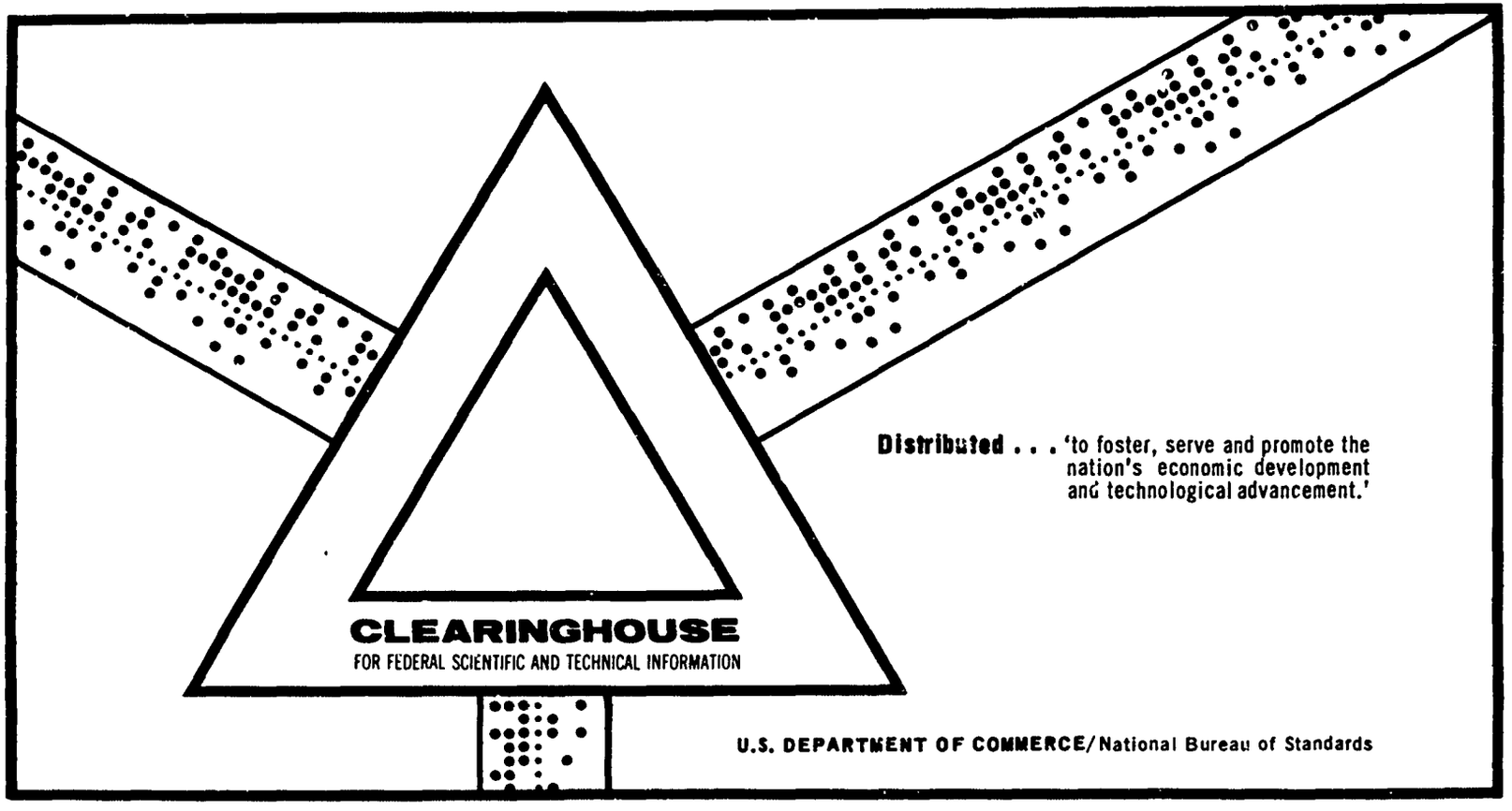
CODED FREQUENCY SHIFT KEYED SEQUENCES WITH APPLICATIONS TO LOW DATA RATE COMMUNICATION AND RADAR

Michael Jon Sites

Stanford University
Stanford, California

September 1969



SU-SEL-69-033

Coded Frequency Shift Keyed Sequences with Applications to Low Data Rate Communication and Radar

by
Michael Jon Sites

September 1969

Technical Report No. 3606-5

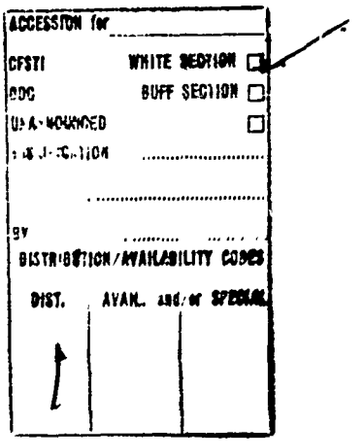
This report was supported in part by the
Joint Services Electronics Programs
U.S. Army, U.S. Navy and U.S. Air Force
Contract Nonr-225(83), NR 373 360

RADIOSCIENCE LABORATORY

STANFORD ELECTRONICS LABORATORIES

STANFORD UNIVERSITY · STANFORD, CALIFORNIA

107


ACCESSION for
CFSTI WHITE SECTION
DDC BUFF SECTION
UNANNOUNCED
JUSTIFICATION
BY
DISTRIBUTION/AVAILABILITY CODES
DIST. AVAIL. and/or SPECIAL

SEL-69-033

CODED FREQUENCY SHIFT KEYED SEQUENCES WITH APPLICATIONS TO LOW DATA RATE COMMUNICATION AND RADAR

by

Michael Jon Sites

September 1969

Technical Report No. 3606-5

This work was supported in part by the
Joint Services Electronics Programs
U.S. Army, U.S. Navy, and U.S. Air Force
under Contract Nonr-225(83), NR 373 360

Radioscience Laboratory
Stanford Electronics Laboratories
Stanford University Stanford, California

ABSTRACT

Coded discrete frequency sequences can provide greatly improved performance over conventional techniques when the fluctuation bandwidth of the communication channel is a significant fraction of the transmission bandwidth. These fluctuations result from medium, equipment, and, in the case of radar, random target variations.

The ability to reliably detect these sequences under frequency shift is investigated using a simple algorithm to calculate an approximation to the true ambiguity function. This investigation leads to certain necessary conditions and a set of coupled equations which permit a sequence to be synthesized from the ambiguity function approximation.

Error rates for CFSK sequences are calculated and criteria for optimizing the performance in rapidly fluctuating channels given. An algorithm for constructing sets of orthogonal sequences with desirable cross-ambiguity properties is developed and the performance of these sets of orthogonal sequences compared with binary code alphabets.

Finally, consideration is given to digital and analog implementations of the special receiver required for the CFSK sequences, error correction coding problems associated with M-ary encoding, and acquisition behavior.

CONTENTS

		Page
I	INTRODUCTION	1
II	DETECTION OF FLUCTUATING SIGNALS	3
	2.1 Introduction	3
	2.2 Receivers for Fluctuating Signals	3
	2.3 The Optimum Detector for Partially Coherent Communication	11
	2.4 Sources and Characterization of Fluctuations	13
	2.4.1 Fluctuations in the Medium	13
	2.4.2 Fluctuations Originating in the Transmitter and Receiver	15
III	CODED FREQUENCY SHIFT KEYED SEQUENCES	
	3.1 Introduction to Coded Frequency Sequences	19
	3.2 CFSK Sequences	19
	3.3 Matched Filter for CFSK Sequences	21
	3.3.1 Coherent and Non-Coherent Matched Filter	21
	3.3.2 Coherent Ambiguity Function	24
	3.3.3 Ambiguity Functions for Non-Coherent Detection	25
	3.3.4 Envelope Ambiguity Function	27
	3.3.5 The Response Lattice	28
	3.4 Calculation of the Response Lattice from the Normalized CFSK Sequence	32
	3.5 Simple Operations on CFSK Sequences and Corresponding Effects on the Response Lattice	34
	3.6 Response Lattices with the "Thumbtack" Property	42
	3.7 Some Necessary Conditions on the Response Lattice for Realizable Sequences	44
IV	MODULATION TECHNIQUES AND ERROR RATES FOR FADING CHANNELS	55
	4.1 Introduction	55
	4.2 Error Rates for Slow Fading--No Diversity	55
	4.3 Diversity Techniques	57
	4.4 Error Rates for Arbitrary Fading Rate and Order of Diversity-Binary Alphabet	60
	4.5 Optimization of Diversity Channel Power and Coded Frequency Sequences	72

CONTENTS (Cont)

		Page
4.6	Selection of Orthogonal Sequences.	77
4.7	M-ary Source Alphabets	79
4.8	Constructing Sets of Orthogonal Sequences.	82
V	RECEIVER IMPLEMENTATION, CODING, AND FREQUENCY ACQUISITION	87
5.1	Implementation of the CFSK Receiver.	87
5.2	Frequency Acquisition of CFSK Sequences.	89
5.3	Coding for Reliable Transmission in M-ary Systems.	90
VI	RECOMMENDATIONS FOR FURTHER INVESTIGATION	94
BIBLIOGRAPHY		95
REFERENCES		98

ILLUSTRATIONS

Figure		Page
2.1	Maximum posterior probability detector for equal energy, equiprobable binary signals.	12
3.1	Frequency shift keyed sequence $\{f_3f_2f_5f_1f_6f_4\}$ with six discrete harmonic frequencies.	20
3.2	Coherent matched filter for $\psi_n(t)$	23
3.3	Non-coherent matched filter for $\psi_n(t)$	23
3.4	Coherent ambiguity function for sequence $\{f_3f_5f_1f_4f_2f_6\}$.	26
3.5	Coherent ambiguity function for sequence $\{f_1f_2f_3f_4f_5f_6\}$.	26
3.6	Non-coherent ambiguity function for sequence $\{f_3f_5f_1f_4f_2f_6\}$	29
3.7	Non-coherent ambiguity functions for sequence $\{f_1f_2f_3f_4f_5f_6\}$	29
3.8	Envelope ambiguity function for sequence $\{f_3f_5f_1f_4f_3f_6\}$.	30
3.9	Envelope ambiguity function for sequence $\{f_1f_2f_3f_4f_5f_6\}$.	30
3.10	Response lattice for sequence $\Gamma = \{325164\}$.	34
3.11	Derivation of the response lattice from the F-matrix $\Gamma = \{325164\}$	36
3.12	Response lattice and F-matrix for eight sequences in {325164} group.	38
3.13	Response lattice for sequence synthesis example.	41
4.1	Probability of error for Rayleigh channel with slow fluctuations	58
4.2	Incoherent matched filter diversity receiver	62
4.3	SNR degradation for exponential fading autocorrelation (n = 1, 10.)	67
4.4a	Probability of error vs total SNR for Rayleigh channel with slow fading M is the number of diversity channels .	69
4.4b	Probability of error vs total SNR for Rayleigh channel with exponential fading autocorrelation (b = 0.10.). . . .	70

ILLUSTRATIONS (Cont)

Figure		Page
4.4c	Probability of error vs total SNR for Rayleigh channel with exponential fading autocorrelation (b = 0.20.). . . .	71
4.5	Probability of error vs diversity channel SNR for energy sharing diversity and diversity channel signal duration proportional to 1/M.	73
5.1	Block diagram of CFSK sequence receiver.	88

ACKNOWLEDGMENT

I wish to express special thanks to Professor Allen M. Peterson for his encouragement and many valuable suggestions. I also wish to thank Professor Bruce B. Lusignan for his help and support.

Financial support for this research was provided by the Communications and Radar Division of Hughes Aircraft Company through Hughes Doctoral Fellowship Program. Computer time and publishing expenses were funded by Tri-Service Contract Nonr-225(83).

Chapter I

INTRODUCTION

Because of the extreme ranges over which planetary radars and deep space probes must operate, narrowband transmissions must be used to obtain a sufficiently large signal-to- noise ratio for reliable detection and communication. However, the propagation medium and equipment instabilities (as well as random target fluctuations in the radar case) may have a fluctuation bandwidth which is a significant fraction of the transmission bandwidth. By utilizing a discrete frequency transmission in which the transmitter frequency is changed in a particular sequence at regular intervals the ratio of transmission to fluctuation bandwidth may be significantly reduced. These coded frequency-shift keyed (CFSK) sequences are identical to discrete frequency codes currently used in large time-bandwidth radars except that the waveform compression is performed after detection rather than before.

In addition to increasing the ratio of transmission to fluctuation bandwidth the codes provide diversity improvement as well as the ability to greatly increase the information rate by M-ary rather than binary encoding.

To provide a basis for later development Chapter II investigates the causes and magnitudes of these fluctuations, the difficulty of estimating and correcting for their effect, and special receivers for fluctuating channels. It is concluded that correction of the effects caused by rapidly fluctuating channels is not feasible.

Next, in Chapter III, in order to study the effects of uncorrected doppler shift and mistuning on the detection of the sequence a simple algorithm for rapidly calculating an approximation to the ambiguity function is derived. In addition, simple necessary conditions and a synthesis procedure for obtaining codes from the ambiguity function approximation is given.

The CSFK sequences are compared in Chapter IV to standard modulation techniques and optimized as a function of available power and fluctuation bandwidth. It is shown that these optimized codes give acceptable error performance in channels where coherent communication techniques would be unusable. The algorithm for the ambiguity function approximation also is useful for examining the mutual interference between sequences.

It is found that a set of orthogonal sequences with low mutual interference can be constructed by means of a simple algorithm and that this set of orthogonal sequences may be used to greatly increase the information rate by M-ary coding.

Chapter V considers digital and analog implementations of the CFSK receiver as well as efficient error control and acquisition behavior.

The most significant results of this investigation are the response lattice concept for representing CFSK sequence ambiguity functions and the resulting algorithm for obtaining the response lattice directly from the sequence, the necessary conditions and procedure for synthesizing CFSK sequences from the response lattice, the conditions for minimizing error rates for CFSK sequences used over fluctuating channels, and the algorithm for synthesizing sets of orthogonal sequences with particular cross-ambiguity properties.

Using these results it is shown that CFSK sequences can provide considerable improvement over conventional techniques when communicating over channels subject to rapid fluctuations.

Chapter II

DETECTION OF FLUCTUATING SIGNALS

2.1 Introduction

In a digital communications system utilizing a nontime-varying channel it is always possible to decrease the bit error probability by increasing the bit length. Although in many real communication links the parameters of the link fluctuate with time, these fluctuations often are negligible for data bandwidths much greater than the rate of fluctuation. However, if the bit duration is increased to the point where the channel fluctuation bandwidth is comparable to the data bandwidth, the error rate does not decrease as rapidly with an increase in bit length as for the time-invariant channel. Ultimately, these time variations, rather than thermal noise, can set a lower limit on the obtainable probability of error. The data rates at which these fluctuations become appreciable will be called low data rates. The region of low data rate is not well defined and depends in detail on the transmitted waveform, the exact form of the channel fluctuations, the probability of error desired, and the receiver implementation. These same considerations occur in detection of radar targets at extreme ranges.

2.2 Receivers for Fluctuating Signals

For purposes of describing the interaction of the channel fluctuations with the transmitted waveform it is convenient to introduce the two-frequency cross-correlation function and its two dimensional Fourier transform, the channel scattering function. The two-frequency cross-correlation function is obtained by transmitting two sinusoidal signals spaced in frequency by Δf. At the receiver one of the sinusoids is delayed relative to the other by a time $\Delta\tau$ and then cross-correlated with the undelayed sinusoid. By varying $\Delta\tau$ and Δf the function can be determined. The correlation function is defined in terms of the complex envelopes of the received scattered waveforms by

$$\mathcal{R}(\Delta\tau,\Delta f) = \mathrm{Re}\left\{\int_{-\infty}^{\infty} \chi(t)\chi^*(t+\Delta\ )\ e^{-j2\pi ft}\ e^{-2\pi(\Delta f+f)(\Delta\tau+t)}\ dt\right\} \qquad (2.1)$$

A section of $\mathcal{R}(\Delta f,\Delta\tau)$ along the $\Delta f = 0$ axis gives the correlation between two points τ sec apart on a transmitted sinusoid. The range of $\Delta\tau$ where $\mathcal{R}(0,\Delta\tau)$ is appreciable will set an upper limit on the interval over which a signal may be coherently integrated in a matched filter. Setting $\Delta\tau = 0$ gives the spaced-frequency correlation function. The range of Δf for which $\mathcal{R}(\Delta f,0)$ is appreciable gives the bandwidth over which the frequency components of a transmission will remain coherent. If the bandwidth of the transmission is too large then the channel will severely distort the received waveform.

The Fourier transform of $\mathcal{R}(\Delta f,\Delta\tau)$ is the channel scattering function $\sigma(\tau,f)$ which is the distribution in time and frequency that the energy of a hypothetical transmitted impulse in time and frequency would have after interacting with the channel. By integrating out τ, the channel power spectrum $\sigma(f)$, given by

$$\sigma(f) = \int_{-\infty}^{\infty} \sigma(\tau,f)\, d\tau \,, \tag{2.2}$$

may be obtained from $\sigma(\tau,f)$. This function represents the spectrum of a continuous sinusoidal transmission after interaction with the channel.

To show explicitly the interaction of the channel, an arbitrary transmitted waveform $u(t)$, and the receiver filter, it is useful to introduce the ambiguity function $\chi_u^2(\tau,f)$ defined as (Price and Green [1960])

$$\chi_u^2(\tau,f) = \left| \int_{-\infty}^{\infty} u(t)\, u(t+\tau) \exp(-j2\pi ft)\, dt \right|^2 \tag{2.3}$$

The function $\chi_u^2(\tau,f)$ gives the square of the magnitude of the response of a matched filter to a frequency translated input. If the filter is not matched to $u(t)$, but has an impulse response $s(t)$, the square of the magnitude of the filter response to $u(t)$, is given by the cross-ambiguity function $\chi_{us}^2(\tau,f)$ defined by

$$\chi_{us}^2(\tau,f) = \left| \int_{-\infty}^{\infty} u(t)\, s(t+\tau) \exp(-j2\pi ft)\, dt \right|^2 \qquad (2.4)$$

The power out of the filter as a function of frequency and time translation for a matched filter is given by

$$P(\tau,f) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \chi_u^2(\tau' - \tau,\ f - f')\, \sigma(\tau',f')\, d\tau'\, df' \ . \qquad (2.5)$$

If the receiver filter is not matched to $u(t)$, $\chi_u^2(\tau,f)$ is replaced by $\chi_{us}^2(\tau,f)$.

The power for a given τ and f displacement is the volume under the product of the ambiguity function (or cross-ambiguity function), shifted by τ and f, and the channel scattering function. This equation suggests that the way to maximize the received power for a particular τ and f is to design χ_u^2 (or χ_{us}^2) to have nearly the same shape as $\sigma(\tau,f)$ in order to maximize the common volume. For instance, if the scattering function is broader in the frequency dimension than the delay, a transmitted waveform which has an ambiguity function broader in frequency offset than in time delay by the same proportion will yield the largest output power. However, this shape matching alone will not guarantee good detectibility.

If $\chi_u^2(\tau,f)$ (or $\chi_{us}^2(\tau,f)$) has appreciable extent along the τ-axis then the envelope and phase of the received waveform will fluctuate over the bit time, because the frequency components of the received waveform are decorrelated by the channel. Conversely, if the scattering function has appreciable extent along the frequency axis, the envelope and phase of the received waveform will fluctuate because segments arriving at different times will be decorrelated.

The rate of fluctuation in relation to the bit length depends on the ratio of the scattering function extent in the frequency domain compared to the extent of the ambiguity function in the time domain. If the ratio is appreciable then the received waveform will be decorrelated within a fraction of a bit time. This circumstance must be avoided if possible because it makes coherent detection difficult.

It will be shown below that despite matching of the ambiguity function shape to that of the scattering function, even the optimum receiver (as defined below) will be degraded if the product of the RMS deviation in delay and frequency of the scattering function approaches or exceeds unity. However, there will still be a bit duration that minimizes the degradation.

Price and Green analyzed the performance of various receiver implementations in terms of their relative probability of detecting the presence of a signal under the condition of an equal false-alarm rate (Neyman-Pearson criterion). They showed that when the received signal and noise are sample functions from independent zero-mean gaussian random processes and the noise is stationary and white with a one-sided spectral density of N_0 W/Hz, the optimum receiver computes a decision variable which is a monotonic function of the likelihood ratio. This decision variable is given by

$$\tilde{D}[\omega_0(t)] = \frac{N_0}{2} \int_{-T_0/2}^{T_0/2} \int_{-T_0/2}^{T_0/2} \omega_0(t)\, \omega_0(t')\, \tilde{F}(t,t')\, dt\, dt' \tag{2.6}$$

where $\omega_0(t)$ is the received signal plus noise and the optimum processing kernel $\tilde{F}(t,t')$ is the solution of the integral equation

$$\int_{-T_0/2}^{T_0/2} \left[\phi_z(t,t') + \frac{N_0}{2}\,\delta(t-t')\right] \tilde{F}(t,t'')\, dt' = \phi_z(t,t'') \;;$$

$$-\frac{T_0}{2} \le t,t' \le \frac{T_0}{2} \tag{2.7}$$

Here ϕ_z is the assumed known correlation function of the part of the received signal representing the channel-distorted transmitted waveform. Whenever the noise density is so great that it dominates the left side of the above equation, it may be solved to give an approximation, $F(t,t')$, to the optimum processing kernel $\tilde{F}(t,t')$:

$$F(t,t') = \frac{2}{N_0} \phi_z(t,t') \simeq \tilde{F}(t,t') \tag{2.8}$$

The decision variable computed by this approximate processor, $D[\omega_0(t)]$, is given by

$$D[\omega_0(t)] = \int_{-T_0/2}^{T_0/2} \int_{-T_0/2}^{T_0/2} \omega_0(t)\, \omega_0(t')\, \phi_z(t,t')\, dt\, dt' \,. \tag{2.9}$$

The loss in detectability of $D[\omega_0(t)]$ compared to the optimum processor $\tilde{D}[\omega_0(t)]$ is bounded by a function of the largest eigenvalue of the correlation function $\phi_z(t,t')$ and the noise density N_0. If λ_0 is the largest solution to the equation

$$\int_{-T_0/2}^{T_0/2} \phi_z(t,t')\, \psi_k(t')\, dt' = \lambda_k \psi_k \tag{2.10}$$

the signal power must be increased by at most the factor $(1 + 2\lambda_0/N_0)$ to give equal detectability, and the condition $2\lambda/N \ll 1$ is sufficient to insure negligible difference in detectability. In addition, if all of the positive eigenvalues of Eq. (2.10) are equal or if λ_0 is the only eigenvalue, then the processor using $\phi_z(t,t')$ has no loss at all compared to the optimum processor, independent of the ratio $2\lambda_0/N_0$.

A special case of the processor defined by Eq. (2-9) is a coherent receiver consisting only of a single linear filter followed by a sampler and squarer where

$$\phi_z(t,t') = g(t)\, g(t') \,. \tag{2.11}$$

The sampler samples the filter output at its maximum value; the squarer does not affect the probability of detection, but allows the coherent receiver to be represented in the same mathematical form as the optimum processor.

By defining the output signal-to-noise ratio as the increase in power of the squared output when the signal is present compared to when the signal is absent divided by the variance of the squared output when the signal is absent and assuming a symmetrical processing kernel, the output signal-to-noise ratio is

$$R = \frac{2}{N_0^2}\left\{\int_{-T_0/2}^{T_0/2}\int_{-T_0/2}^{T_0/2}[\phi_z(t,t')\ K(t,t')\ dt\,dt']^2\right/\left.\int_{-T_0/2}^{T_0/2}\int_{-T_0/2}^{T_0/2}K^2(t,t')\ dt\,dt'\right\} \tag{2.12}$$

It is interesting to compare the output SNR of coherent filter receiver and the receiver using the kernel $\phi_z(t,t')$. By Eq. (2.12) the SNR of the coherent filter receiver is

$$R_F = \frac{2}{N_0^2}\left[\int_{-T_0/2}^{T_0/2}\int_{-T_0/2}^{T_0/2}\phi_z(t,t')\ g(t)\ g(t')\ dt\,dt'\right]^2\Bigg/\int_{-T_0/2}^{T_0/2}\int_{-T_0/2}^{T_0/2}g(t)\ g(t')\ dt\,dt' \tag{2.13}$$

Using the Rayleigh-Ritz inequality, R_F has an upper bound given by $R_F < 2(\lambda_0/N_0)^2$, with equality occurring if $g(t) = \psi_0(t)$, where $\psi_0(t)$ is the eigenfunction associated with λ_0. For the more general SNR maximizing processor using $K(t,t')$ application of the Schwartz inequality shows that the SNR is upper bounded by

$$R \le \frac{2}{N_0^2}\int_{-T_0/2}^{T_0/2}\int_{-T_0/2}^{T_0/2}\phi_z^2(t,t')\ dt\ dt' = \frac{2}{N_0^2}\sum_{k=0}^{\infty}\lambda_k^2 = R_{max} \tag{2.14}$$

with equality whenever $K(t,t')$ is proportional to $\phi_z(t,t')$. Thus the ratio of the maximum SNR for coherent filter processing to the maximum SNR for a processing kernel equal to the correlation function $\phi_z(t,t')$ is

$$R_{max}/R_{F,max} = \sum_{k=0}^{\infty} \lambda_k^2/\lambda_0^2 = 1 + \sum_{k=1}^{\infty} \lambda_k^2/\lambda_0^2 \geq 1 \qquad (2.15)$$

Thus, unless λ_0 is unique, the output SNR of the coherent filter receiver will always be smaller than that of the processor with a kernel $\phi_z(t,t')$. Also, as long as $2\lambda_0/N_0$ is much less than unity, output SNR is a good indication of the relative ability of processors to detect a signal [Price, 1965].

The following hierarchy of processor output SNR shows the relative performance of coherent filter processors, correlation function matched processors and the ultimate upper bound on SNR, $2(\bar{E}/N_0)^2$:

$$R_{F,max} = 2\left(\frac{\lambda_0}{N_0}\right)^2 \leq R_{D,max} = 2\sum_{k=0}^{\infty}\left(\frac{\lambda_k}{N_0}\right)^2 \leq 2\left(\sum_{k=0}^{\infty}\frac{\lambda_k}{n_0}\right)^2 = 2\left(\frac{\bar{E}}{N_0}\right)^2 . \qquad (2.16)$$

From the above expression, the ultimate upper bound on SNR of $2(\bar{E}/N_0)^2$ can be obtained with a coherent processor only if λ_0 and its associated eigenfunction are unique. In this case the two receivers are equivalent. If there are two or more eigenvalues, then the available signal energy is distributed among the multiple eigenfunctions, causing the coherent-filter squarer to be inferior to the processor matched to the correlation function, which, in turn, no longer achieves the upper bound of $2(\bar{E}/N_0)^2$ for SNR.

It is instructive to consider the eigenfunctions for a particular channel autocorrelation function given by

$$R(\Delta t,\Delta f) = r_0 e^{-2\pi B|\Delta t|} \qquad (2.17)$$

(where B is the fading bandwidth and r_0 is the average power) which corresponds to a scattering function

$$\sigma(\tau,f) = [(Br_0/\pi)/(B^2 + f^2)]\ \delta(\tau) \qquad (2.18)$$

The eigenvalues of this correlation function, for $BT \ll 1$, are

$$\lambda_0 \simeq \left[1 - \frac{2\pi BT}{3}\right] r_0 T = \left[1 - \frac{BT}{3}\right] \bar{E} \tag{2.19}$$

and,

$$\lambda_k \simeq \left[4\pi BT/k^2\pi^2\right] r_0 T , \quad k = 1, 2, \ldots \tag{2.20}$$

Since this is a lowpass autocorrelation function, to calculate the eigenvalues for a fluctuating phase, narrowband signal which has two orthogonal eigenfunctions associated with each eigenvalue, the eigenvalues calculated above must be divided by a factor of 2. Therefore, for a channel with an exponential fading spectrum, the largest eigenvalue, for small values of BT, is given by

$$\lambda_0' = \lambda_0/2 = \frac{1}{2}\left\{1 + (2\pi)\frac{BT}{3}\right\} r_0 T . \tag{2.21}$$

Since the output SNR for the coherent filter processor, R_F, is bounded by

$$SNR_0 \leq 2(\lambda_0/N_0)^2 \simeq \frac{1}{2}\left[1 - (2\pi)\frac{BT}{3}\right]^2 \left[\frac{r_0 T}{N_0}\right]^2 = \frac{1}{2}\left[1 - (2\pi)\frac{BT}{3}\right]^2 \left(\frac{E}{N_0}\right)^2$$

$$\leq \frac{1}{2}(\bar{E}_z/N_0)^2 \tag{2.22}$$

where $\bar{E}_z$ is the average received signal energy, the ratio of output SNR for the two detectors is

$$\frac{R_p}{R_F} = \frac{2\sum(\lambda_k/N_0)^2}{2(\lambda_0/N_0)^2} = 1 + \frac{1}{\lambda_0^2}\sum_{k=1}^{\infty}\lambda_k^2 = 1 + \sum_{k=1}^{\infty}\left[1 - BT/3\right]^{-2}\left[2BT/k^2\pi^2\right]^2 , \quad BT \ll 1 \tag{2.23}$$

The above expression shows that for non-zero BT the SNR optimizing processor will do better than the matched filter. For small values of BT (BT $\ll$ 1), however, there will be little degradation in SNR from using the matched filter processor.

2.3 The Optimum Detector for Partially Coherent Communication

A communications link operating over a time-varying channel must estimate the phase of the received signal in the presence of both additive and (for a channel with negligible multipath) multiplicative noise. If a phase lock loop is used to derive an estimate of the signal phase from the carrier, pilot tone, or virtual carrier, the estimate will differ from the true phase because of the noise, causing a detection loss. The optimum receiver must account for the variance of the phase error. Viterbi [1966] gives the exact probability density for the phase error, ϕ, of a first order phase-lock-loop as

$$p(\phi) = \exp(\alpha \cos \phi)/2\pi I_0(\alpha) \qquad (2.24)$$

where α is the signal-to-noise ratio in the loop bandwidth. The optimum detector for equiprobable, equal energy signals (in a maximum posterior-probability sense) is shown in Fig. (2.1), where y_{is} and y_{ic} are estimates of the in-phase and quadrature components of the two possible signals $S_i, i = 0,1$. This detector reduces to a coherent detector when α is large and to an envelope correlation detector when α is small. For a nontime-varying channel, α depends only on thermal noise. However, if the carrier is frequency modulated as a result of multiplicative fading or oscillator instability then the effective loop SNR, α', will be decreased to $\alpha' \simeq \left[\alpha^{-2} + \sigma_f^2\right]^{-1/2}$, where σ_f^2 is the mean-square error caused by the inability of a narrow loop to track the rapidly (compared to the loop bandwidth) varying carrier. Generally, σ_f will be a decreasing function of the loop bandwidth, whereas α^{-1} will increase with the loop bandwidth. Thus, for a given amount of frequency spreading and additive noise density, there is an optimum loop bandwidth that will minimize the total mean-square phase error. However, if this mean-square phase error is on the order of 0.25 rad^2 or greater, the optimum detector

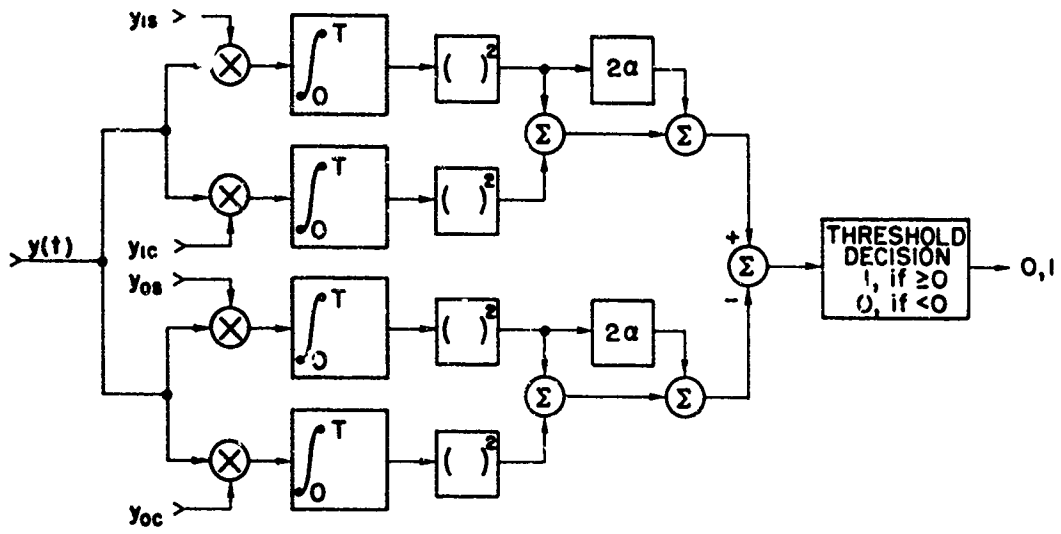


Fig. 2.1. MAXIMUM POSTERIOR PROBABILITY DETECTOR FOR EQUAL ENERGY, EQUIPROBABLE BINARY SIGNALS. (Adapted from Viterbi [1966].)

will be closely approximated by an envelope correlation detector. The situation where both frequency instability and thermal noise are simultaneously present contrasts with the usual circumstance where the data bandwidth is much greater than the fluctuation bandwidth of the carrier or pilot tone. In the latter circumstance the bandwidth of the phase-lock-loop can be sufficiently narrow that a good phase reference is obtained. Conversely, if the fluctuation bandwidth is on the order of the data bandwidth, it will be impossible to obtain a good phase reference. Viterbi showed that because of the inability to obtain good phase estimates, for binary probabilities of error on the order of 10^{-2} to 10^{-5}, incoherent FSK is superior to coherent PSK for values of α' less than about 5.

Since α' is proportional to the total RMS phase error, phase fluctuations originating in the transmission medium and short-term oscillator instability can severely degrade receiver performance even when the thermal noise is relatively low.

2.4 Sources and Characterization of Fluctuations

The fluctuations experienced by a signal can be attributed to one of two causes: fluctuations originating in the medium and fluctuations originating in the transmitting and receiving equipment.

2.4.1 Fluctuations in the Medium

An electromagnetic wave at a frequency much greater than the critical frequency propagating through an electron gas requires an amount of time τ, where (Evans and Hagfors [1968], p. 113)

$$\tau = \frac{1}{c} \int_0^R \left(1 - \frac{Nc^2}{\pi m f^2}\right)^{-1/2} dr \ . \tag{2.25}$$

If $Ne^2/\pi m f^2 \ll 1$, then, using the binomial theorem,

$$\tau \approx \frac{1}{c} \int_0^R \left[1 + \frac{Ne^2}{2\pi m f^2}\right] dr \ . \tag{2.26}$$

The increase in propagation time over the free space value is

$$\Delta\tau \approx \frac{1}{2c} \int_0^R \frac{Ne^2}{\pi m f^2} dr \ . \tag{2.27}$$

As long as the total integrated electron density of the propagation path is not a function of time the received signal will not fluctuate. However, if the electron density does change the variation in delay will be observed as doppler shifting of the received signal. Since, for a delay increase of $\Delta\tau$ the phase path has been increased by an amount $\Delta\varphi$, where

$$\Delta\varphi \approx f\Delta\tau \approx \frac{1}{2c} \int_0^R \frac{Ne^2}{\pi m f} dr \ , \tag{2.28}$$

the doppler shift Δf of the received signal will be

$$\Delta f \approx -\frac{d}{dt}(\Delta\varphi) = \frac{-e^2}{2\pi cmf}\frac{d}{dt}\int_0^R N\,dr \quad . \tag{2.29}$$

If the integrated electron density changes irregularly the received signal will be broadened in frequency. This broadening may be represented as an increase in the frequency extent of the channel scattering function.

Usually electron density fluctuations can be neglected, but when the phase path is very long and the data bandwidth is small, as it would be in the case of an interplanetary probe or a planetary lander, even relatively small density variations can have an appreciable effect on error rates. Whenever a number of independently varying irregularities exist along the phase path the RMS variation is the square root of the sum of the mean-squared frequency variation contributed by each irregularity, i.e.,

$$\Delta f_{rms} = \left[\overline{(\Delta f)^2}\right]^{1/2} = \frac{e^2}{2\pi mcf}\left\{\sum_{\ell=1}^{n}\overline{\left[\int_{r_\ell}^{r_\ell+\Delta_\ell}\left(\frac{dN}{dt}\right)dr\right]^2}\right\}^{1/2} \tag{2.30}$$

where r_ℓ is the distance to the "ℓ"th irregularity, Δ_ℓ is the effective size, and n is the number of independently fluctuating irregularities. Assuming the phase path may be represented by n contiguous fluctuation cells all with effective length $\Delta = \Delta_\ell$ and identical fluctuation statistics, Eq. (2.30) can be written

$$\Delta f_{rms} = \frac{e^2}{2\pi mcf}\sqrt{R/\Delta}\left\{\overline{\left[\int_0^{\Delta}\left(\frac{dN}{dt}\right)dr\right]^2}\right\}^{1/2} \tag{2.31}$$

If N is approximately constant over a fluctuation "cell," then

$$\Delta f_{rms} \approx \frac{e^2}{2\pi mcf}\sqrt{R\Delta}\left\{\overline{\left(\frac{dN}{dt}\right)^2}\right\}^{1/2} \quad . \tag{2.32}$$

The expression in Eq. (2.32) applies to the effects of the earth's ionosphere as well as those of the interplanetary plasma. For transmitters close to the earth variations originating in the ionosphere will be the major contributors to frequency spreading, but over distances many times the earth's radius the tenuous interplanetary plasma can also have an appreciable contribution.

If a probe should pass behind the sun the solar corona can have a major effect. Shapiro [1964] estimated the increase in time delay for a raypath passing near the sun at a distance r solar radii to be

$$\Delta\tau \approx \frac{1.47 \times 10^{14}}{f^2 r} \text{ sec} . \tag{2.33}$$

Then, the rms frequency spreading is approximated by

$$\Delta f_{rms} \approx f\left\{\left[\frac{d}{dt}(\Delta\tau)\right]^2\right\}^{1/2} \tag{2.34}$$

The turbulence of the solar corona suggests that $\Delta\tau$ could vary quite rapidly. Assuming that $\Delta\tau$ can vary as much as 0.1 percent in a second,

$$\Delta f_{rms} \approx \frac{1.47 \times 10^{11}}{fr} \text{ Hz} . \tag{2.35}$$

Even for a frequency of several GHz and paths where $r \leq 10$ the frequency spreading could exceed 5 Hz. This amount of spreading would be catastrophic for a system with a data bandwidth of only a few cycles.

Frequency spreads of this magnitude were reported for Mariner IV paths which passed through the corona [Goldstein, et al, 1967].

2.4.2 Fluctuations Originating in the Transmitter and Receiver

A major source of short-term instabilities which originate in equipment is lack of short-term stability of oscillators used for frequency

generation and translation. If an oscillator spectrum is examined with a high resolution spectrum analyzer, the spectrum consists of a narrow peak which decreases rapidly as $1/f$ or $1/f^2$ as regions separated by f Hz from the nominal center frequency are examined until a plateau of uniform noise is reached. The $1/f^2$ (or $1/f$) dependency is the result of phase modulation by noise originating in the oscillator itself, such as $1/f$ transistor noise and thermal noise. Power supply ripple and external signals coupling into the oscillator circuitry can also produce phase modulation, although these sources usually produce modulation components at specific frequencies. The finite-width spectrum of the nonideal oscillator means that over a time $1/\Delta f$, where Δf is the half-power line width, the oscillator phase will wander approximately one radian. If $1/\Delta f$ is comparable to (or greater than) the bit length, and this noisy oscillator is used to frequency translate an otherwise nonfluctuating signal, coherent addition in the data filter of signal energy representing different parts of the pulse will not occur, and the signal-to-noise ratio from a matched filter detector will be degraded unless phase estimation and correction is performed prior to detection. As has been previously discussed, when the fluctuation spectrum is on the order of the data bandwidth, any phase estimate derived from the noisy, rapidly fluctuating signal may be so poor that adequate phase correction is impossible.

The spectral purity of oscillators used for frequency generation and translation in spacecraft can be improved by phase-locking to a highly stable carrier transmitted by the tracking station; however, at low data rates both the medium and instabilities of the phase-lock-loop VCO will limit the accuracy of the phase-locking. Fluctuations in the medium cause frequency spreading of the stable carrier; any frequency source phase-locked to this received carrier will have the same frequency fluctuations. Even if the medium were perfectly stable and the effects of thermal noise could be neglected it would still be impossible to achieve perfect phase-lock because of noise in the voltage controlled oscillator being locked to the stable carrier. This noise causes phase modulation of the VCO spectrum; however, as long as the bandwidth of the tracking loop is large compared to the width of the VCO spectrum, the phase noise will be nulled

out by the feedback and the VCO spectrum will be pure. The mean-square loop phase error which results from the VCO phase noise is

$$\overline{\theta_L^2} = \frac{1}{2\pi} \int_0^\infty \gamma(\omega) \left|1 - H(\omega)\right|^2 d\omega, \qquad (2.36)$$

where $\gamma(\omega)$ is the one-sided spectral density of the oscillator phase noise in $(\text{rad})^2/\text{Hz}$ and $H(\omega)$ is the closed loop transfer function. If the carrier strength is so weak that a narrow loop bandwidth is required to exclude thermal noise, then it can happen that an appreciable portion of the VCO spectrum extends beyond the loop bandwidth. In this case only the low frequency components of the phase modulation are removed and the high frequency components cause the derived carrier phase to rapidly fluctuate. Even when small these fluctuations can degrade coherent detection and, if large enough, can cause a tracking loop to skip cycles. Whenever this is the case either the loop bandwidth must be increased or the VCO stability must be improved.

Unfortunately, increasing loop bandwidth increases the phase error of thermal noise origin and improving the VCO stability is often not feasible because of power and weight considerations or tracking range requirements (which might mean that a VCXO could not be used).

One more consideration which aggravates the problem of achieving a pure spectrum is that frequency multiplication, often required to generate stable frequencies in the microwave region, causes sideband phase noise power to increase as the square of the multiplication factor. The resulting phase noise increase can cause an oscillator which might be satisfactory at a lower frequency to become unsuitable when multiplied.

In addition to instabilities which originate in oscillators, another source of instability is the path lergth variations caused by mechanical antenna vibration. For the microwave frequencies used for space communication even relatively small mechanical movements can cause significant frequency spreading, particularly if they are rapid. If the movement is $b \sin \omega_a t$, then the frequency spreading is

$$\Delta\omega = \frac{d}{dt}(\Delta\varphi) = \frac{d}{dt}\left(\frac{b}{\lambda} 2\pi \sin \omega_a t\right) = \frac{2\pi b\omega_a}{\lambda} \cos(\omega_a t) \quad . \qquad (2.37)$$

Assuming a vibrational movement to wavelength ratio of 0.01 and a vibrational frequency of 1 Hz, the rms frequency deviation is

$$\Delta f_{rms} = \frac{\sqrt{2}\pi b\omega_a}{\lambda} = 0.28 \text{ Hz} \quad .$$

For data bandwidths on the order of 1 Hz this amount of frequency spreading would be significant.

The value of b/λ postulated would, for $\lambda = 10$ cm, correspond to a movement of only 1 mm. Motions of this magnitude could occur in antennas carried by planetary landers if buffeted by high winds or in large receiving antennas on the earth.

Thus for data bandwidths on the order of 1 Hz phase fluctuations of various origins such as the medium, oscillator instability, thermal noise, and mechanical vibration can combine to produce enough frequency spreading that phase coherence over a bit length is not, and cannot, be obtained. For this reason techniques using shorter pulses and incoherent detection, even though sub-optimal in more stable communication links, become attractive.

Chapter III

CODED FREQUENCY SHIFT KEYED SEQUENCES

3.1 Introduction to Coded Frequency Sequences

In Chapter II a number of sources of frequency spreading with magnitudes on the order of 0.01 Hz to 1 Hz were identified and their degrading effects investigated. A primary cause of degradation was found to be the lack of coherence over a bit length. This observation suggests that if the length of the bit could be made shorter, coherence could be obtained. However, if the transmitter is power limited, decreasing the bit length will reduce the received energy per bit and hence the detectibility. One way out of this dilemma is to sample the envelope of the received signal at intervals approximately equal to the coherence time (assumed to be less than the bit length). This technique will give approximately independent samples of signal plus noise which can be added to arrive at the decision variable. In this technique the receiver essentially becomes a weighted radiometer.

These independent envelope samples could just as well have been obtained by stepping the transmitter at intervals between M frequencies whose spacing is a multiple of the reciprocal of the bit duration. An example of such a frequency shift keyed sequence is shown in Fig. 3.1. As long as the subpulse duration is short compared to the coherence time the signal energy will be approximately coherent over the subpulse.

3.2 CFSK Sequences

Systems using binary phase or frequency shift keying are common in both radar and communications. In these systems a binary sequence modulates the carrier amplitude, phase, and frequency in discrete increments. In general, the transmitted waveform produced by discrete modulation can be represented in complex notation (Cook and Bernfield [1967])

$$\psi(t) = \sum_{i=1}^{L} \alpha_i E_i(t) \exp[j(\omega_0 + \omega_i)\, t + \theta_i)] \qquad (3.1)$$

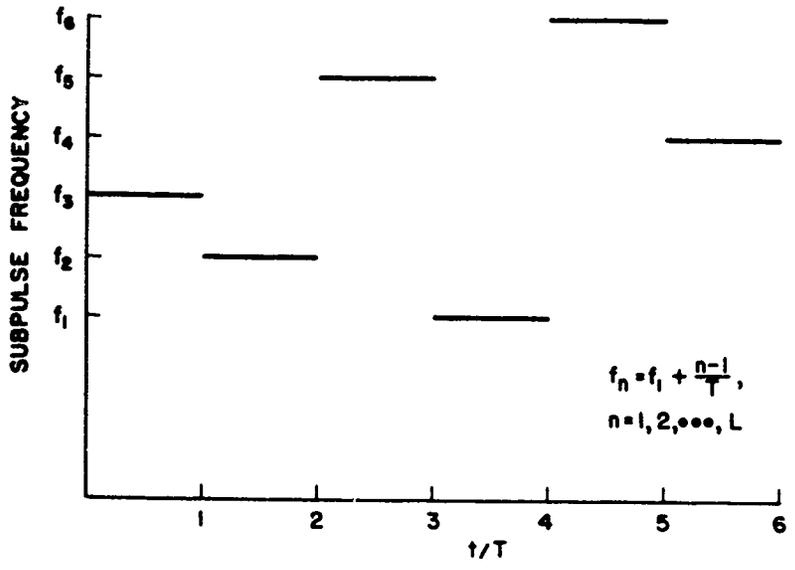


Fig. 3.1. FREQUENCY SHIFT KEYED SEQUENCE $\{f_3 f_2 f_5 f_1 f_6 f_4\}$ WITH SIX DISCRETE HARMONIC FREQUENCIES.

where

$$E_i(t) = 1, \ (i - 1)\ \delta \leq t \leq i\delta$$

$$= 0, \ t > i \ \text{ or } \ t < (i - 1)\ \delta$$

and α_i, ω_i, and θ_i are the amplitude, radian frequency, and phase, respectively, associated with the envelope $E_i(t)$. For a binary PSK communication system, α_i is constant, $\omega_i = 0$, and $\theta_i = \pm\pi/2$ depending on whether a 1 or 0 was transmitted. For a binary FSK system θ_i and α_i are constant and $\omega_i = \omega_1$ or ω_2, representing a transmitted 1 or 0. These binary systems can be generalized by allowing the phase or frequency to have M discrete values instead of only two. Such an M-ary communication system can transmit $\log_2 M$ bits/symbol instead of the one bit/symbol of a binary system. In an M-ary FSK system reception of ω_i is taken to mean that the message symbol S_i corresponding to ω_i was transmitted, since each symbol is associated with a unique frequency.

There is also the possibility of associating a unique sequence Ω_n of transmitted frequencies with each message symbol, where the sequence

$$\Omega_n = \{\omega_{n1}\omega_{n2} \cdots \omega_{nL}\} \tag{3.2}$$

represents a frequency shift keyed sequence with radian frequencies ω_{ni}. if there are L available transmitter frequencies they can be ordered without replacement into $L!$ sequences. Each of these $L!$ sequences can be associated with a message symbol S_i such that

$$S_1 : \psi_1(t) = \sum_{i=1}^{L} \alpha_i E_i(t) \exp[j\{(\omega_0 + \omega_{1i})\ t + \theta_i\}]$$

$$S_2 : \psi_2(t) = \sum_{i=1}^{L} \alpha_i E_i(t) \exp[j\{(\omega_0 + \omega_{2i})\ t + \theta_i\}]$$

$$\vdots$$

$$S_n : \psi_n(t) = \sum_{i=1}^{L} \alpha_i E_i(t) \exp[j\{(2_0 + \omega_{ni})\ t + \theta_i\}] \tag{3.3}$$

where α_i and θ_i are constants and ω_{ni} is the ith element of the frequency sequence Ω_n.

3.3 Matched Filter for CFSK Sequences

3.3.1 Coherent and Non-coherent Matched Filter

It is well known that for a non-spread channel the signal-to-noise maximizing processor is a filter matched to the transmitted waveform in

the sense that, if $\psi_n(t)$ is the transmitted waveform, the impulse response of the filter $h(t)$ is

$$h(t) = \psi_n^*(-t) \tag{3.4}$$

For the frequency sequences of Section 2.2 the matched filter may be implemented as a parallel bank of filters, each matched to a particular subpulse $b_i(t)$ of $\psi_n(t)$, where

$$b_i(t) = P(t) \exp[j[(\omega_0 + \omega_i)\, t + \theta_i]] , \tag{3.5}$$

and

$$P(t) = \begin{cases} 1 & 0 \leq t < \delta \\ 0 & \text{otherwise} . \end{cases}$$

Each matched filter is followed by a time delay to bring the subpulses into synchronism. The time delay will depend on the particular frequency sequence Ω_n. If ω_{nk} is the kth element of Ω_n then the filter matched to $b_i(t)$ must be followed by a delay of $(L - k + 1)\, \delta = \tau_{ni}$, where L is the number of frequencies in the sequence. These delays are necessary to permit coherent addition of the subpulses. After coherent addition, the resulting waveform is detected to yield the decision variable for $\psi_n(t)$. Figure 3.2 illustrates one possible implementation of a coherent matched filter $\psi_n(t)$. Implementation of the coherent matched filter requires that the time delays be correct to within a small fraction of the carrier period to permit coherent summation. Fluctuation of the channel will adversely affect the subpulse-to-subpulse coherence, with the result that coherent summation is not feasible. Under these circumstances non-coherent summation obtained by following each delay with an envelope squarer and summing the envelopes will be necessary. Figure 3.3 shows the implementation for non-coherent detection of $\psi_n(t)$. The two implementations are not equivalent: envelope summation reduces the CFSK sequence time-bandwidth product from L^2 obtained with coherent summation to just L.

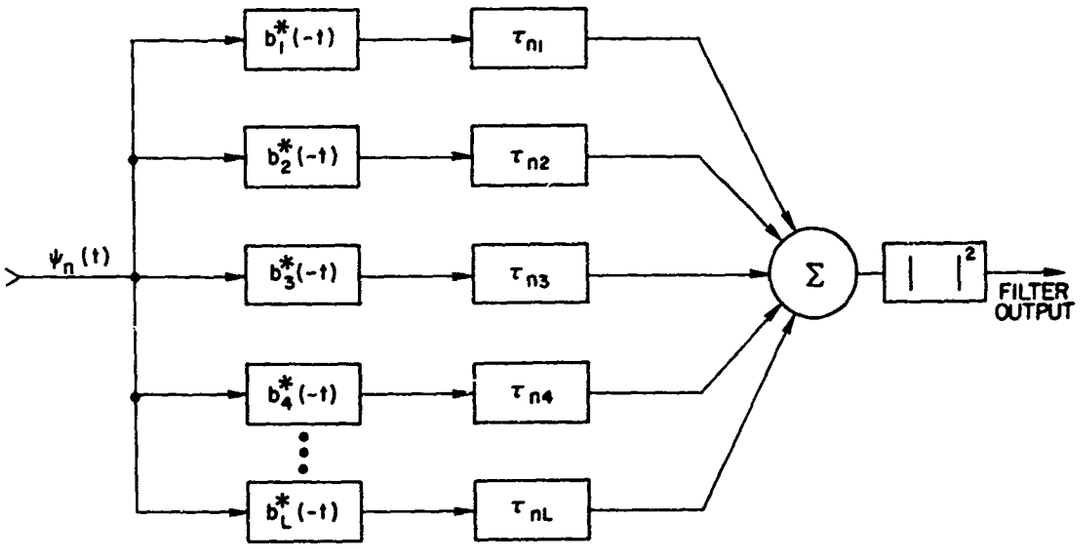


Fig. 3.2. COHERENT MATCHED FILTER FOR $\psi_n(t)$.

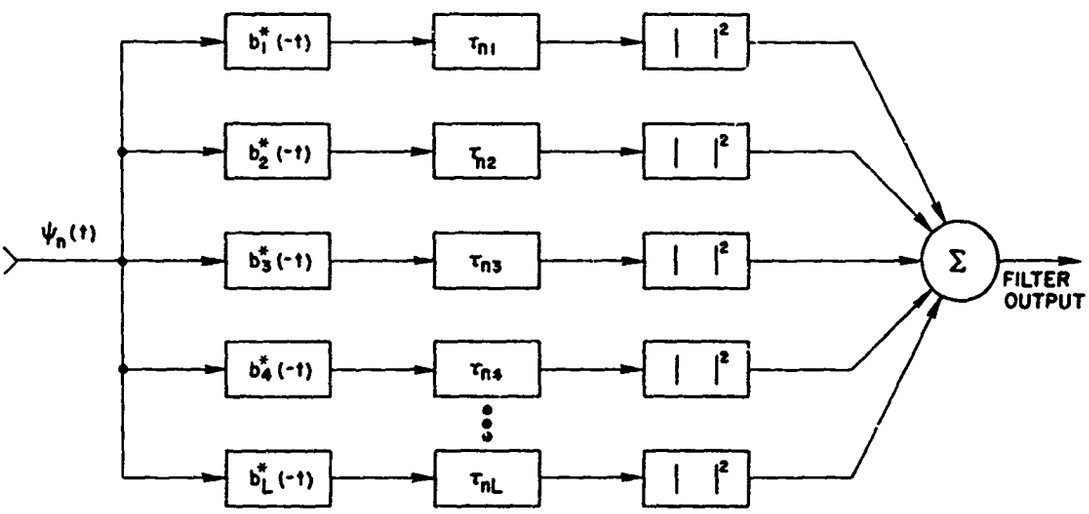


Fig. 3.3. NON-COHERENT MATCHED FILTER FOR $\psi_n(t)$.

The processors of Figs. 3.2 and 3.3 can also be implemented as taps on a single delay line.

3.3.2 Coherent Ambiguity Function

The coherent ambiguity function $|\chi_n(\tau,\phi)|$ represents the envelope as a function of delay of a matched filter output when the waveform to which to which the filter is matched is displaced in frequency by an amount ϕ.

If the terms in $\psi_n(t)$ representing the frequency modulation are included in the envelope term so that Eq. (3.1) becomes

$$\chi_n(t) = \sum_{i=1}^{L} [\alpha_i E_i(t) \exp\{j(\omega_i t) + \theta_i\}] \exp(j\omega_0 t)$$

$$= \exp(j\omega_0 t) \sum_{i=1}^{L} u_i(t) \ , \tag{3.6}$$

then the ambiguity function for $\psi_n(t)$ is

$$|\chi_{\psi_n}(\tau,\phi)| = \left| \int_{-\infty}^{\infty} \psi_n(t) \ \psi_n^*(t + \tau) \exp(2\pi\phi t) \ dt \right|$$

$$= \left| \exp(-j\omega_0\tau) \int_{-\infty}^{\infty} \sum_{i=1}^{L} u_i(t) \sum_{m=1}^{L} u_m^*(t + \tau) \exp(2\pi\phi t) \ dt \right| . \tag{3.7}$$

Taking the envelope of $\chi_{\psi_n}(\tau,\phi)$ eliminates the fine structure contributed by the carrier frequency term so that Eq. (3.7) becomes

$$\left| \chi_{\psi_n}(\tau,\phi) \right| = \left| \int_{-\infty}^{\infty} \sum_{i=1}^{L} \sum_{m=1}^{L} u_i(t) \ u_m^*(t + \tau) \exp(2\pi\phi t) \ dt \right| . \tag{3.8}$$

Since the response to an input $u_i(t)$ of a filter matched to a particular subpulse $u_m(t)$ as a function of frequency offset and delay is

$$\int_{-\infty}^{\infty} u_i(t)\, u_m^*(t+\tau)\, \exp(2\pi\phi t)\, dt = \chi_{u,n}(\tau - \Delta\tau_{im},\ \phi - \Delta\phi_{im,n}) \qquad (3.9)$$

where $\Delta\tau_{im} = (i - m)\,\delta$ and $\Delta\phi_{im,n} = \frac{1}{2\pi}(\omega_{ni} - \omega_{nm})$, the coherent ambiguity function for $\psi_n(t)$ can be expressed as a double sum of time and frequency shifted replicas of the function $\chi_{u,n}(\tau,\phi)$ (where the envelope is obtained after summation):

$$\left|\chi_{\psi_n}(\tau,\phi)\right| = \left|\sum_{i=1}^{L}\sum_{m=1}^{L} \chi_{u,n}(\tau - \Delta\tau_{im},\ \phi - \Delta\phi_{im,n})\right| . \qquad (3.10)$$

Figures 3.4 and 3.5 give the coherent ambiguity function for the sequences $\{f_3 f_5 f_1 f_4 f_2 f_6\}$ and $\{f_1 f_2 f_3 f_4 f_5 f_6\}$, respectively, where

$$f_n = f_1 + \frac{n-1}{T} .$$

The response shown in Fig. 3.4 is an example of a "thumbtack" ambiguity function (see Section 3.6).

3.3.3 Ambiguity Functions for Non-coherent Detection

If the filter outputs are envelope detected before being summed, then the ambiguity function for $\chi_n(t)$ no longer has the form of Eq. (3.10). Under these circumstances the envelope of the response of the mth filter to $\psi_n(t)$ is

$$\chi_{u_{m,n}}(\tau,\phi) = \left|\int_{-\infty}^{\infty} \sum_{i=1}^{L} u_i(t)\, u_m^*(t+t)\, \exp(2\pi\phi t)\, dt\right|$$

$$= \left|\sum_{i=1}^{L} \chi_{u,n}(\tau - \Delta\tau_{im},\ \phi - \Delta\phi_{im,n})\right| \qquad (3.11)$$

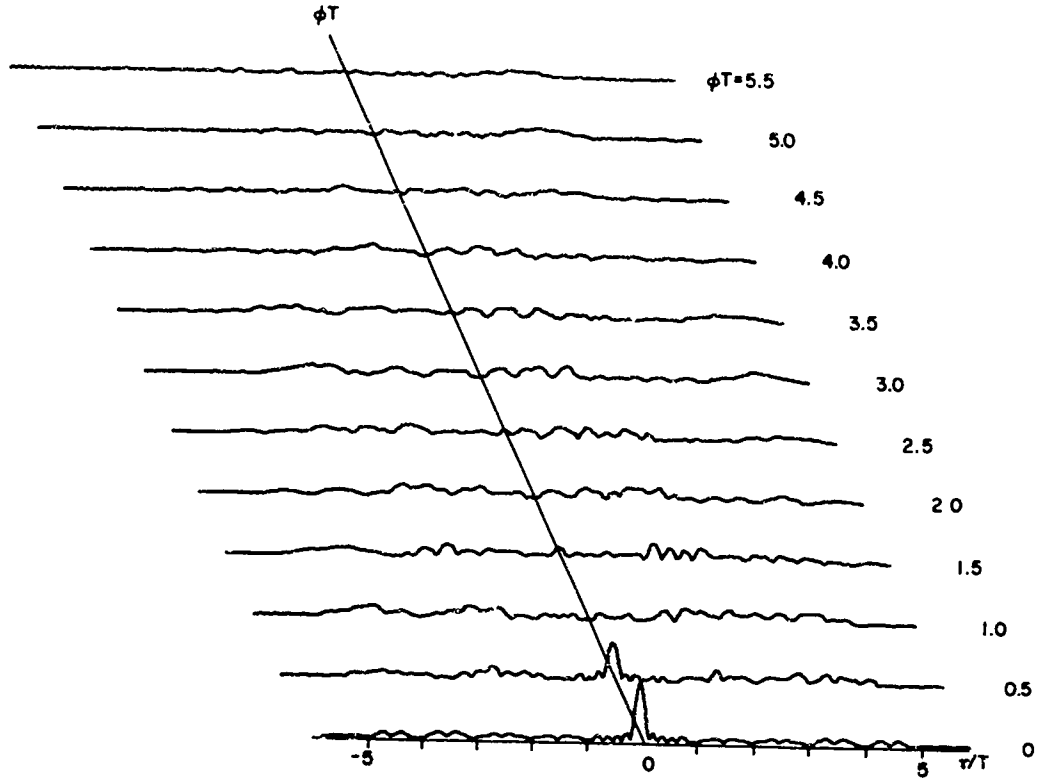


Fig. 3.4. COHERENT AMBIGUITY FUNCTION FOR SEQUENCE $\{f_3 f_5 f_1 f_4 f_2 f_6\}$.

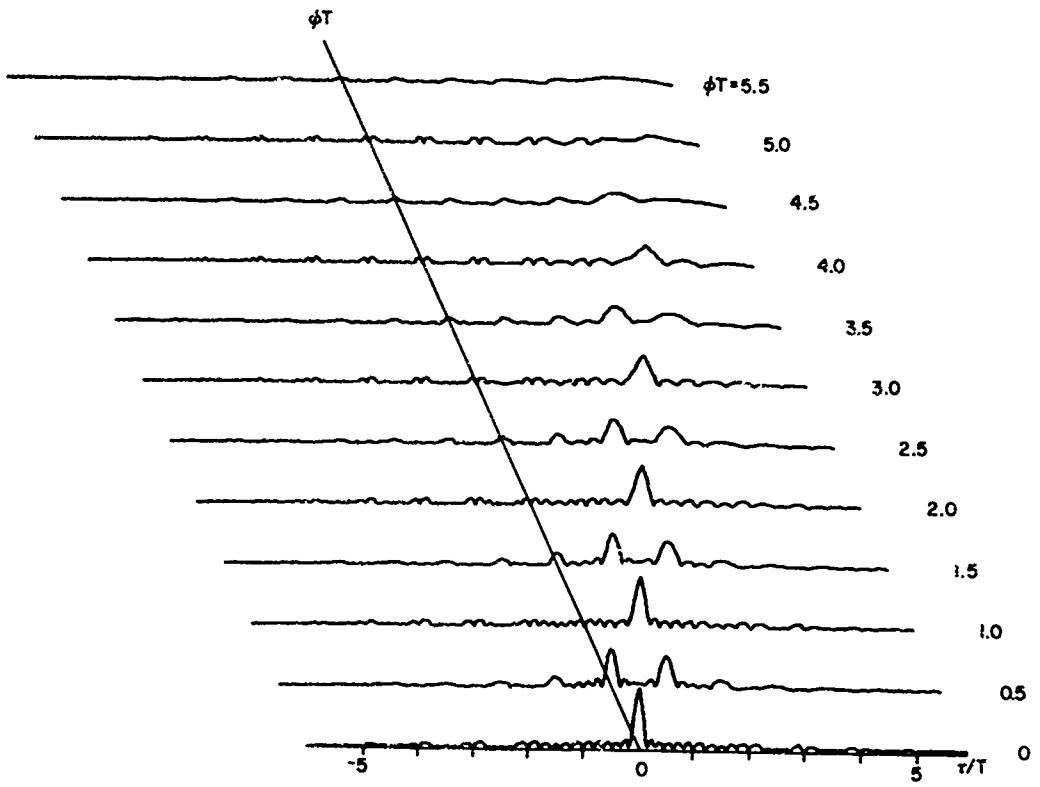


Fig. 3.5. COHERENT AMBIGUITY FUNCTION FOR SEQUENCE $\{f_1 f_2 f_3 f_4 f_5 f_6\}$.

Summation of these envelopes gives the non-coherent ambiguity function for the sequence $\psi_n(t)$:

$$\chi'_{\psi_n}(\tau,\phi) = \sum_{m=1}^{L} \chi_{u_{m,n}}(\tau,\phi)$$

$$= \sum_{m=1}^{L} \left| \sum_{i=1}^{L} \chi_{u,n}(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \right| . \qquad (3.12)$$

Since $|\chi'_{\psi_n}(\tau,\phi)| = \chi'_{\psi_n}(\tau,\phi)$, the non-coherent ambiguity function $|\chi'_{\psi_n}(\tau,\phi)|$ is:

$$\left|\chi'_{\psi_n}(\tau,\phi)\right| = \sum_{m=1}^{L} \left| \sum_{i=1}^{L} \chi_{u,n}(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \right| \qquad (3.13)$$

3.3.4 Envelope Ambiguity Function

Applying the inequality

$$\left| \sum_{L=1}^{n} a_i \right| \leq \sum_{L=1}^{n} \left| a_i \right| \qquad (3.14)$$

to Eq. (3.13) for the non-coherent ambiguity function produces an upper bound:

$$\chi'_{\psi_n}(\tau,\phi) = \sum_{m=1}^{L} \left| \sum_{i=1}^{L} \chi_{u,n}(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \right|$$

$$\leq \sum_{m=1}^{L} \sum_{i=1}^{L} \left| \chi_{u,n}(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \right| \qquad (3.15)$$

This upper bound will be defined as the envelope ambiguity function $\tilde{\chi}_{\psi_n}(\tau,\phi)$, associated with the waveform $\psi_n(t)$. The envelope ambiguity function is obtained by summing the subpulse ambiguity functions shifted in delay and frequency.

Since, by application of the inequality in Eq. (3.14) the non-coherent ambiguity function upper bounds the coherent ambiguity function, the following sequence of inequalities results:

$$|\chi_{\psi_n}(\tau,\phi)| = \left| \sum_{m=1}^{L} \sum_{i=1}^{L} \chi_{u,n}(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \right|$$

$$\leq \sum_{m=1}^{L} \left| \sum_{i=1}^{L} \chi_{u,n}(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \right| = |\chi'_{\psi_n}(\tau,\phi)|$$

$$\leq \sum_{m=1}^{L} \sum_{i=1}^{L} \left| \chi_{u,n}(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \right| = \left| \tilde{\chi}_{\psi_n}(\tau,\phi) \right| \qquad (3.16)$$

As long as there is negligible overlap of the time and frequency shifted ambiguity functions, the envelope ambiguity function will be a tight upper bound on the non-coherent ambiguity function. Figures 3.6 through 3.9 are the calculated non-coherent and envelope ambiguity function for two different CFSK sequences. Because of the non-coherent summation, the central peak is broader than for coherent summation.

3.3.5 The Response Lattice

The video ambiguity function can be written as a two-dimensional convolution of the subpulse ambiguity function envelope with a two-dimensional delta function lattice by expanding the envelope ambiguity function as

$$\tilde{\chi}_{\psi_n}(\tau,\phi) = |\chi_{u,n}(\tau,\phi)| \; ** \sum_{m=1}^{L} \sum_{m=1}^{L} \delta(\tau - \Delta\tau_{im}, \phi - \Delta\phi_{im,n}) \; . \qquad (3.17)$$

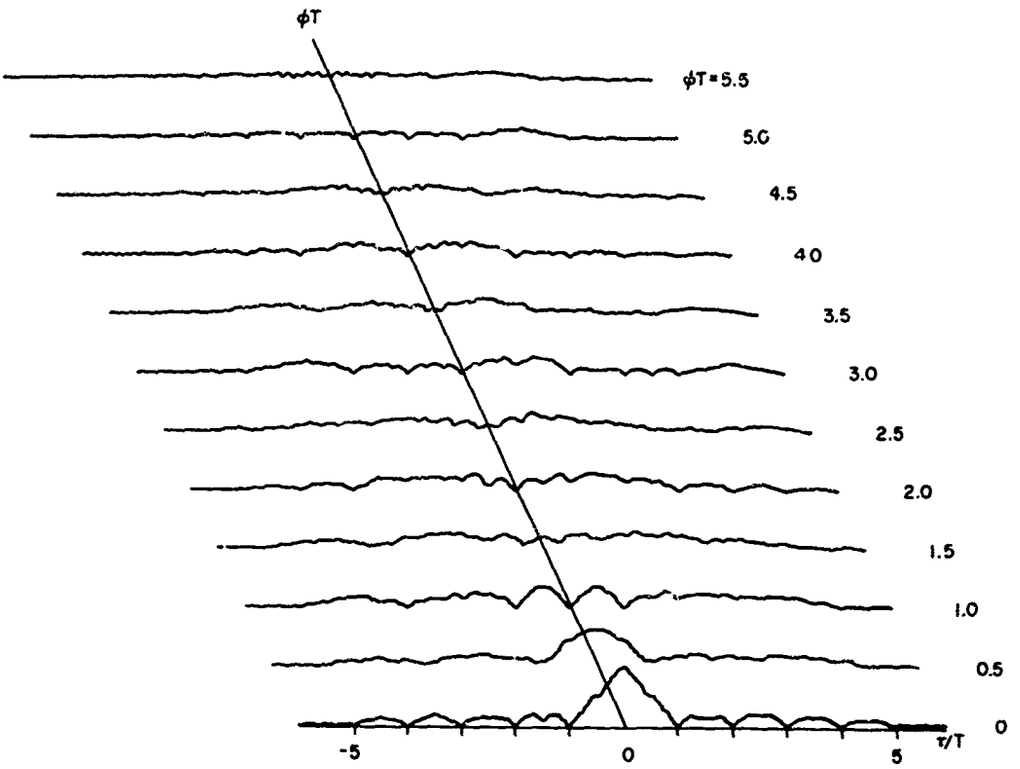


Fig. 3.6. NON-COHERENT AMBIGUITY FUNCTION FOR SEQUENCE $\{f_3 f_5 f_1 f_4 f_2 f_6\}$.

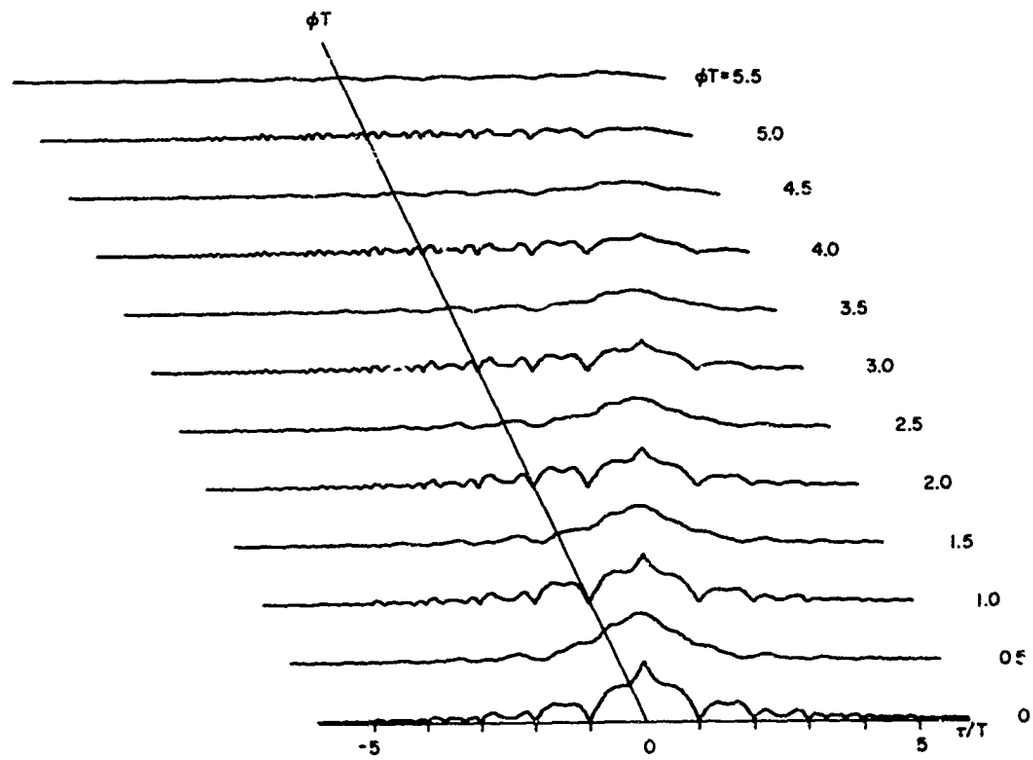


Fig. 3.7. NON-COHERENT AMBIGUITY FUNCTIONS FOR SEQUENCE $\{f_1 f_2 f_3 f_4 f_5 f_6\}$.

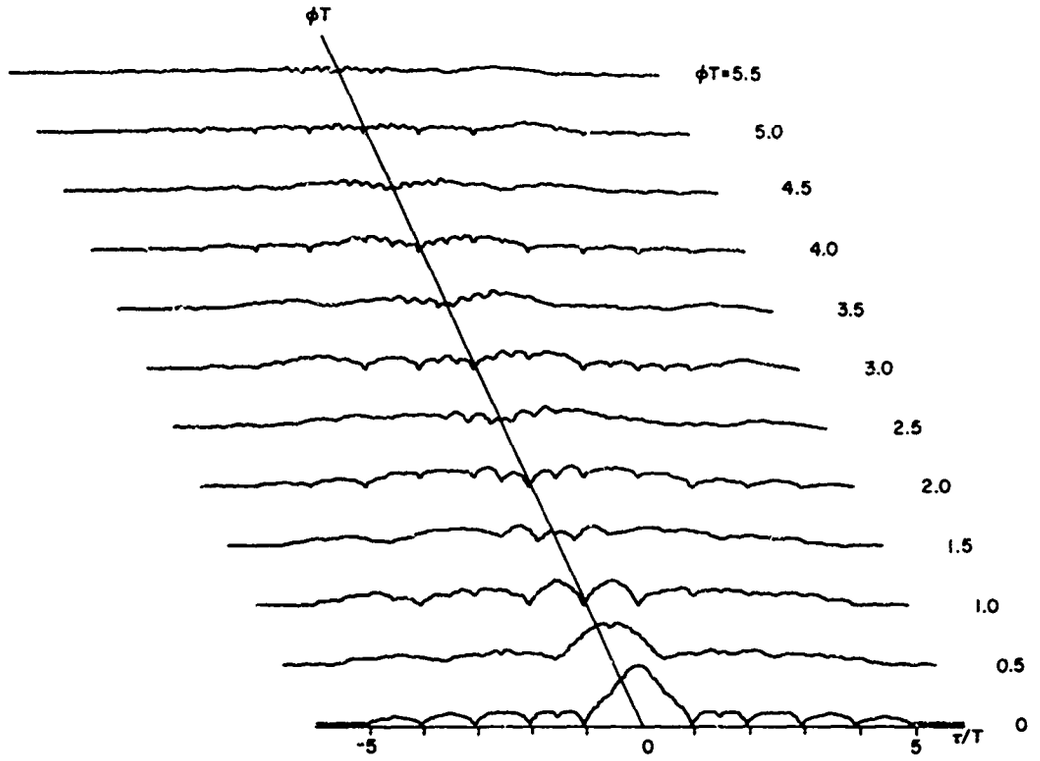


Fig. 3.8. ENVELOPE AMBIGUITY FUNCTION FOR SEQUENCE $\{f_3 f_5 f_1 f_4 f_3 f_6\}$.

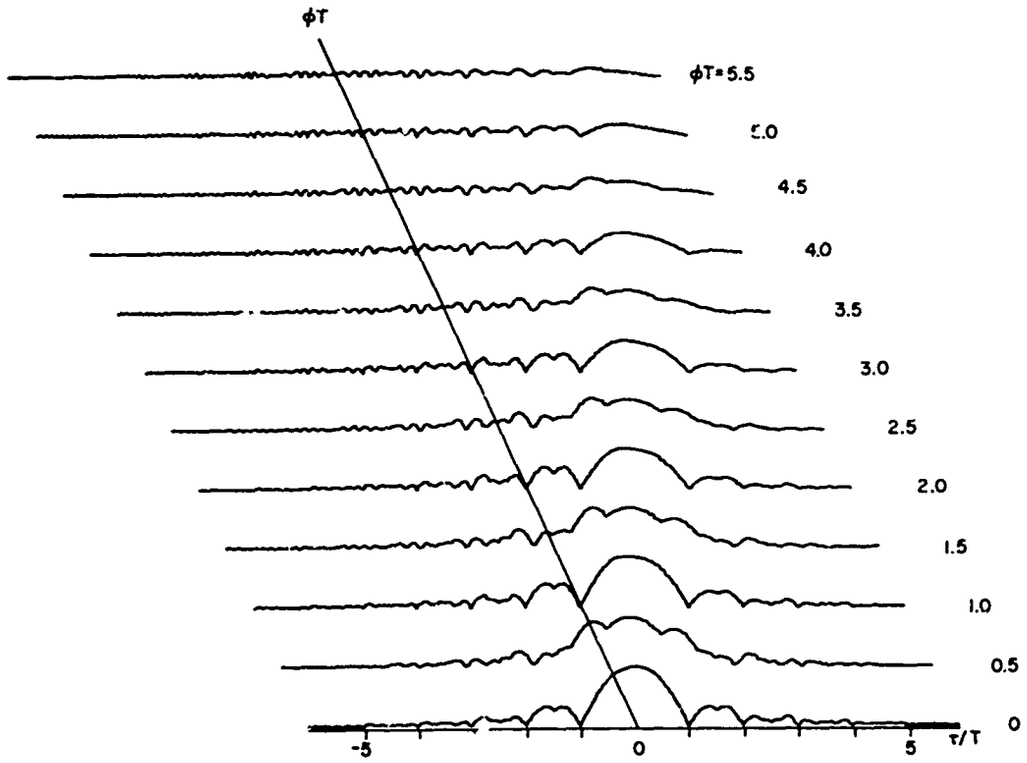


Fig. 3.9. ENVELOPE AMBIGUITY FUNCTION FOR SEQUENCE $\{f_1 f_2 f_3 f_4 f_5 f_6\}$.

The term containing the double summation of delta functions is defined to be the response lattice. Since $\tilde{\chi}_{\psi_n}(\tau,\phi)$ is a sum of real-valued terms, it also is real.

For the uniform pulse trains considered so far,

$$\Delta\tau_{im} = (i - m)\ \delta \tag{3.18}$$

and

$$\Delta\phi_{im,n} = \frac{1}{2\pi}\ (\omega_{ni} - \omega_{nm})\ . \tag{3.19}$$

By normalizing with respect to the subpulse length δ, Eq. (3.17) can be written as

$$\tilde{\chi}_{\psi_n}\left(\frac{\tau}{\delta},\phi\,o\right) = \left|\chi_{u,n}\left(\frac{\tau}{\delta},\phi\,\delta\right)\right| \ **\ \sum_{m=1}^{L}\sum_{i=1}^{L}\ \delta\left[\frac{\tau}{\delta} - (i-m)\ ,\ \phi\delta - (\gamma_{ni} - \gamma_{nm})\right]. \tag{3.20}$$

where

$$\gamma_{ni} = (\omega_{ni}\delta)\ \frac{1}{2\pi}\ .$$

If $\omega_{ni+1} - \omega_{ni} = \frac{2\pi}{\delta}\ P$ where P is an integer, then

$$\Gamma_n = (\gamma_{n1}\ \gamma_{n2}\ \ldots\ \gamma_{nL}) \tag{3.21}$$

will be a sequence of integers representing the frequency modulation normalized with respect to $1/\delta$. If the ω_{ni} are further restricted to the band of frequencies

$$\omega_0 + \frac{2\pi}{\delta} \le \omega_{ni} \le \left(\frac{2\pi}{\delta}\right) L + \omega_0\ ,$$

then Γ_n will be an ordered sequence of the integers $\{1,2,3,\ldots,L\}$.

Changing the sequence order will change the response lattice and the transmitter waveform. However, because of the relationship between the ambiguity function (coherent or non-coherent) and the response lattice, inspection of the effects of sequence order rearrangement on the response lattice will permit the new ambiguity function to be quickly bounded. The facility with which the response lattice can be calculated can be calculated will permit very rapid examination of the gross characteristics of the ambiguity function, particularly regions in the τ,ϕ plane where isolated peaks of ambiguity can occur. After promising sequences are found using the response lattice, the coherent or non-coherent ambiguity function can be calculated in detail.

The response lattice also will be useful aid in the synthesis of desirable ambiguity functions.

3.4 Calculation of the Response Lattice from the Normalized CFSK Sequence

This section develops a simple algorithm for computing the response lattice directly from the normalized CFSK sequence, Γ_n. This algorithm will also provide insight for the derivation of certain fundamental properties of the response lattice which restrict the distribution of ambiguity in the delay-doppler plane.

If Ω_n is shifted upward by an amount $\Delta\omega = \frac{2\pi k}{\delta}$, where k is an integer, then $\Gamma^{(k)}$, the normalized frequency shifted sequence is given by

$$\Gamma_n^{(k)} = \{\gamma_1 + k \;\; \gamma_2 + k \;\; \ldots \;\; \gamma_m + k\} \; . \tag{3.22}$$

Thus, under frequency shift some of the elements of $\Gamma_n^{(k)}$ will appear identical to some of the elements of Γ_n. Since the normalized delay between two elements of Γ_n, γ_s and γ_{n+s}, such that $\gamma_{n+s} - \gamma_s = k$, is simply equal to n, a normalized frequency shift of k will result in the signal energy corresponding to γ_s appearing at the output of the matched filter delayed by an amount $n\delta$.

This relationship between the sequence of frequencies Ω_n and the time-frequency normalized sequence Γ_n suggests a simple algorithm for

computing the response lattice directly from Γ_n. Let d_{ij} represent the normalized delay n between $\gamma_k = j$ and some element γ_{k+n} such that $\gamma_{k+n} = i + j$ (i assumed positive). Defining the distance between all possible pairs such that i is positive as above and arranging the d_{ij} in an array where the ith row of the array represents the normalized delays between all elements with a normalized frequency spacing of i gives the information required to construct the response lattice by inspection. This is done by interpreting the value of each d_{ij} as the number of units of δ from the zero-delay axis and i as the number of units of $1/\delta$ from the zero-doppler axis where the energy of a single frequency shifted (by an amount i/δ) subpulse will be found. Thus, there is a simple, direct relationship between the normalized frequency sequence and the response lattice, which in turn is directly related via a two-dimensional convolution to the distribution of ambiguity in the delay-doppler plane.

For example, the sequence 325164 has a d_{ij} array given by

$$\{d_{ij}\}_{325164} = \begin{matrix} -2 & -1 & 5 & -3 & 2 \\ -3 & 4 & 2 & -1 & \\ 2 & 1 & 4 & & \\ -1 & 3 & & & \\ 1 & & & & \end{matrix} \qquad (3.23)$$

Examination of this array shows that for a doppler shift of $1/\delta$ there will be delta functions at delays of -2δ, $-\delta$, 5δ, -3δ, and 2δ located $1/\delta$ from the zero-doppler axis. Similarly for a doppler shift of $2/\delta$ there will be delta functions at delays of -3δ, 4δ, 2δ, and $-\delta$ located $2/\delta$ from the zero-doppler axis. The remainder of the off-doppler responses are found in the same manner. Figure 3.10 shows the complete response lattice for this sequence. Since the response lattice is skew-symmetric, only the upper-half-plane, corresponding to positive doppler shifts, need be calculated.

At this point, it is not at all obvious what effect reordering a particular sequence will have on the response lattice without actually

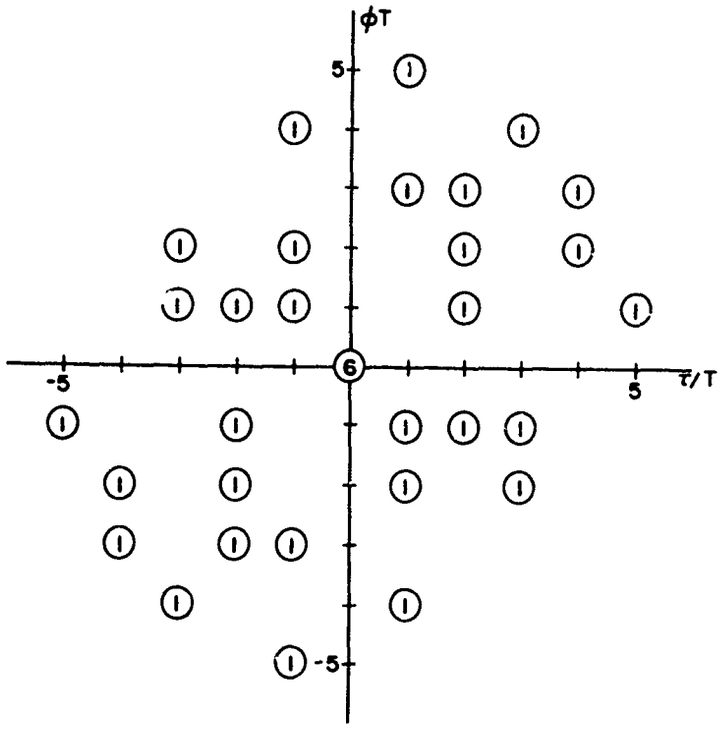


Fig. 3.10. RESPONSE LATTICE FOR SEQUENCE $\Gamma = \{325164\}$.

calculating the d_{ij} array. The next section will develop a simple relationship between operations on a sequence and operations on the response lattice.

3.5 Simple Operations on CFSK Sequences and Corresponding Effects on the Response Lattice

The response lattice may be derived directly from the d_{ij} array as was done in Section 3.4, or may be obtained by the alternate method outlined below.

As the first step in this alternate calculation the sequence Γ_n is transformed into a unitary $M \times M$ matrix F where the Kth column is related to γ_K by

$$f_{jK} = \begin{cases} 1, & \text{if and only if } j = M + 1 - s, \text{ where } \gamma_K = s \\ 0, & \text{otherwise} \end{cases} \tag{3.24}$$

The F matrix can be interpreted as a two-dimensional representation in time and frequency of the normalized CFSK sequence. Thus, for the sequence

$$\Gamma_n = \{\gamma_1\ \gamma_2\ \ldots\ \gamma_6\} = \{3\ 2\ 5\ 1\ 6\ 4\}\ ,$$

the corresponding F matrix is

$$F_{325164} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix} \qquad (3.25)$$

By defining the F matrix corresponding to $\Gamma_n^{(K)}$, the frequency shifted sequence, as $F^{(K)}$ where $F^{(K)}$ is the F matrix with the last $K(K \leq M-1)$ rows removed, it follows that

$$F^{(1)} = \begin{bmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{bmatrix} \qquad (3.26)$$

The remainder of the $F^{(K)}$ are similarly defined. If these $F^{(K)}$ are now considered to just be ordered arrays in the delay-doppler plane and if each of these arrays are shifted to the left or right until the 1's in the last remaining row are aligned with the 1 in the last row of the F array, then the number of ones occupying each cell of the resulting array will be the strength of the corresponding impulse of the response lattice, as shown in Fig. 3.11.

F - matrix :

$$\begin{matrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \underline{1} & 0 & 0 \end{matrix}$$

$F^{(1)}$ - matrix:

$$\begin{matrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & \underline{1} & 0 & 0 & 0 & 0 \end{matrix}$$

$F^{(2)}$ - matrix:

$$\begin{matrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ \underline{1} & 0 & 0 & 0 & 0 & 0 \end{matrix}$$

$F^{(3)}$ - matrix:

$$\begin{matrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \underline{1} \end{matrix}$$

$F^{(4)}$ - matrix :

$$\begin{matrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \underline{1} & 0 & 0 & 0 \end{matrix}$$

$F^{(5)}$ - matrix:

$$\begin{matrix} 0 & 0 & 0 & 0 & 1 & 0 \end{matrix}$$

Note: Underlined numbers represent responses at the origin.

Fig. 3.11. DERIVATION OF THE RESPONSE LATTICE FROM THE F-MATRIX $\Gamma = \{325164\}$.

Obtaining the response lattice in this way brings out an interesting feature of the F matrix: any rotation and/or time-reversal of the F matrix will result in a corresponding rotation and/or time-reversal of

the response lattice. Since the F matrix uniquely defines a sequence, each of these operations on the F matrix yields a new (but not necessarily unique) sequence. In general, up to eight unique sequences may be derived from the eight possible rotations and/or mirror images of the F matrix. These eight sequences, the response lattice, and the F matrices are shown for the sequence {325164} in Fig. 3.12.

From Fig. 3.12 it can be seen that every sequence in the group can be derived from any one initial sequence by a simple rotation and/or transposition of the F matrix. This relationship is shown in detail by Table 3.1 for the sequence {325164}.

It is also possible to obtain the other sequences in a particular group by operating directly on the sequences without having to go through the operations on the F matrix.

First, the time-reverse sequence corresponding to a given sequence will be defined as the sequence obtained by reversing to element order of the original sequence, and the complementary sequence will be the (M+1)'s complement of the elements in the sequence. In addition to these two elementary operations, a new sequence may also be constructed from the delays given each frequency; this sequence will be called the delay sequence. Of course, all three or any combination of these operations can be performed sequentially on a given sequence. Thus, there are seven additional sequences that can be obtained by performing or not performing each of the above operations. Table 3.2 illustrates this interrelationship between the sequences and the elementary operations. Table 3.2 is similar to 3.1 except that the former is the transformation to be performed on the F-matrix corresponding to a particular sequence, while Table 3.2 gives the operations to be performed on the sequences themselves. By comparing Tables 3.1 and 3.2 it can be seen that each transformation of the F-matrix is representable by corresponding operations on the sequence. Reference to Fig. 3.13, which shows the eight possible orientations of the F-matrix and the associated effects on the orientation of the response lattice, gives the information necessary to relate operations on the sequence, transformations on the F-matrix, and the corresponding effects on the response lattice of these operations. These relationships are summarized

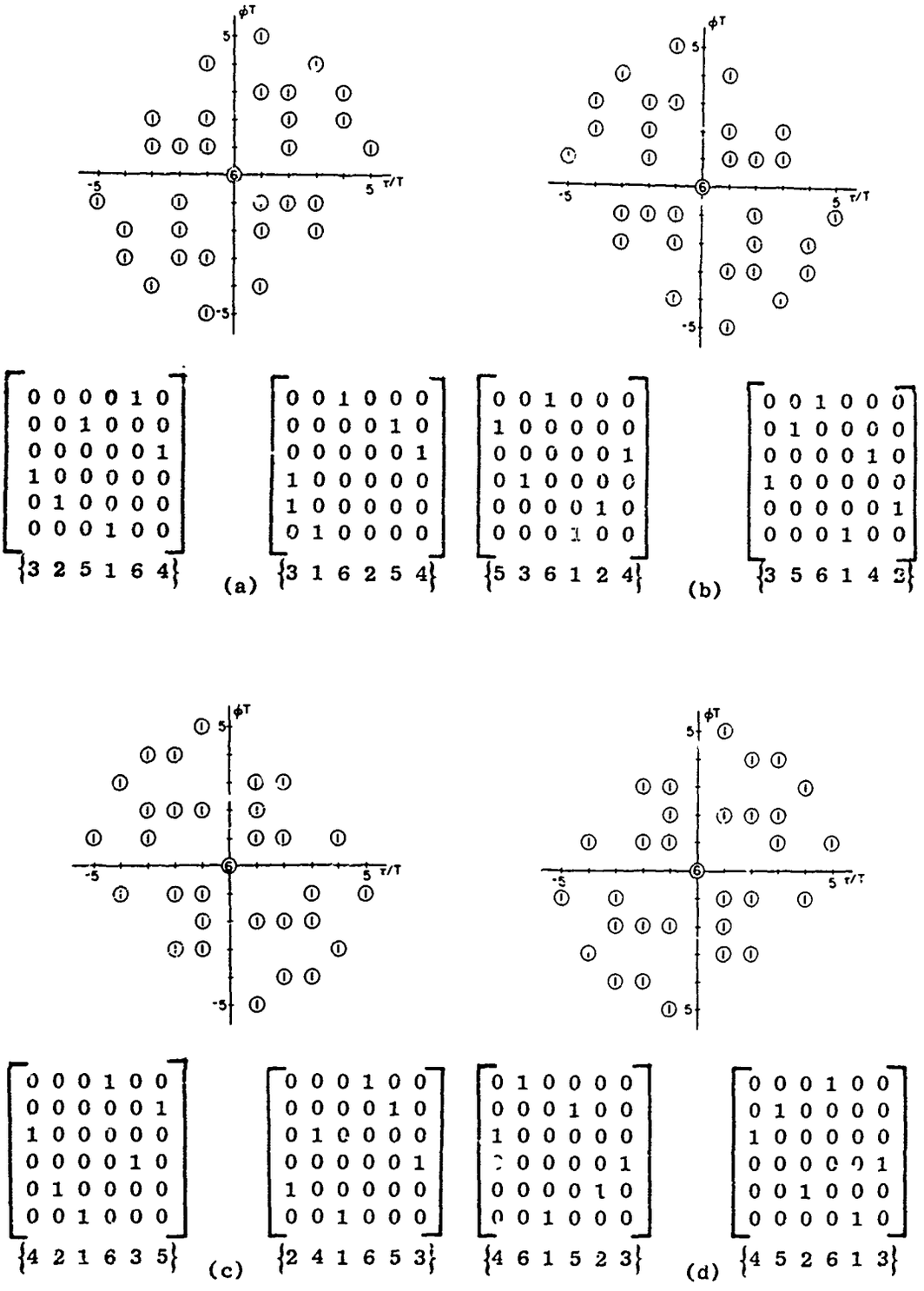


Fig. 3.12. RESPONSE LATTICE AND F-MATRIX FOR EIGHT SEQUENCES IN {325164} GROUP.

Table 3.1

EXAMPLE OF OPERATIONS OF F-MATRIX CORRESPONDING TO SEQUENCE TRANSFORMATION

	325164	421635	241653	316254	461523	356142	536124	452613
325164	no op.	LT	RT	LT, RT	+90°, RT	-90°	+90°	+90°, LT
421635	LT	no op.	LT, RT	RT	+90°	+90°, LT	+90°, RT	-90°
241653	RT	LT, RT	no op.	LT	-90°	+90°, RT	+90°, LT	+90°
316254	LT, RT	RT	LT	no op.	+90°, LT	+90°	-90°	+90°, RT
461523	+90°, RT	+90°	-90°	+90°, LT	no op.	LT	RT	LT, RT
356142	-90°	+90°, LT	+90°, RT	+90°	LT	no op.	LT, RT	RT
536124	+90°	+90°, RT	+90°, LT	-90°	RT	LT, RT	no op.	LT
452613	+90°, LT	-90°	+90°	+90°, RT	LT, RT	RT	LT	no op.

where: +90° = 90° c.c.w. rotation of F-matrix
-90° = c.w. rotation of F-matrix
RT = right transpose of F-matrix
LT = left transpose of F-matrix
no op. = no operation on F-matrix

Table 3.2

	325164	421635	241653	316254	461523	356142	536124	452613
325164	no op.	τ	τ, TR, C	TR, C	TR	τ, C	τ, TR	C
421635	τ	no op.	TR, C	τ, TR, C	τ, TR	C	TR	τ, C
241653	τ, TR, C	TR, C	no op.	τ	τ, C	TR	C	τ, TR
316254	TR, C	τ, TR, C	τ	no op.	C	τ, TR	τ, C	TR
461523	TR	τ, TR	τ, C	C	no op.	τ	τ, TR, C	TR, C
356142	τ, C	C	TR	τ, TR	τ	no op.	TR, C	τ, TR, C
536124	τ, TR	TR	C	τ, C	τ, TR, C	TR, C	no op.	τ
452613	C	τ, C	τ, TR	TR	TR, C	τ, TR, C	τ	no op.

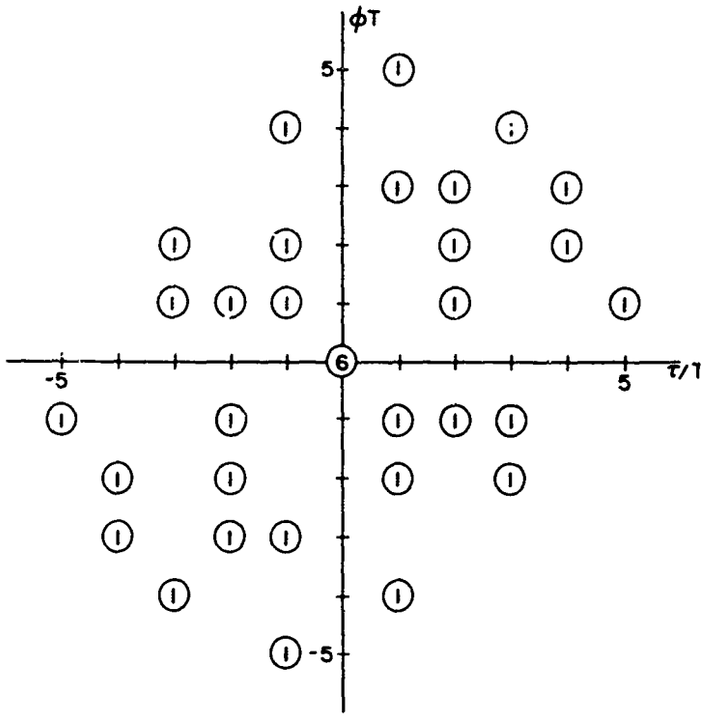


Fig. 3.13. RESPONSE LATTICE FOR SEQUENCE SYNTHESIS EXAMPLE.

in Table 3.3. Several interesting properties are obvious from Table 3.3: first, there are only four distinct response lattices associated with the eight possible sequences; and second, two sequences which are the time-reverse complements of each other have identical response lattices. Furthermore, the operations on the F-matrix which result in identical response lattices also fall into four pairs of opposite operations: the left-transpose, right-transpose pair; the ±90° rotation pair, the 90° rotation followed by a left- or right-transpose pair, and the no-operation, left-transpose-followed-by-a-right-transpose pair.

Table 3.3

CORRESPONDENCE BETWEEN OPERATIONS ON THE F-MATRIX, THE SEQUENCE, AND THE RESPONSE LATTICE

Operation on F-Matrix	Operation on Sequence	Effect on Response Lattice
LT	τ	time reversal, 90° rotation
RT	τ, TR, C	time reversal, 90° rotation
LT, RT	TR, C	no effect
+90°	τ, TR	90° rotation
-90°	τ, C	90° rotation
+90°, RT	TR	time reversal
+90°, LT	C	time reversal
no operation	no operation	no effect

The properties permit the problem of manipulating the response lattice to be transformed into the much simpler problem of operation on either the F-matrix or the sequence.

In the next section a class of sequences with a particular desirable property will be analyzed using these techniques.

3.6 Response Lattices with the "Thumbtack" Property

In Section 3.5 analysis of relationship between manipulation of the F-matrix and the corresponding effects on the response lattice showed that every sequence has up to seven companion sequences which may be easily derived from the first sequence. In this section these relationships will be used to classify into groups all sequences of length 6 which have a "thumbtack" response lattice.

A "thumbtack" response lattice has the property that for a sequence of length M the impulse at $\tau = 0, \phi = 0$ has strength M and every

other impulse located elsewhere in the $\tau - \phi$ plane has a strength of exactly one. Such a response lattice would be useful in a radar system where both delay and frequency shift must be simultaneously estimated or in a communication system where estimation of doppler shift is required to permit frequency tracking.

The "thumbtack" sequences are interesting because there is a simple necessary and sufficient condition for a particular sequence to have a "thumbtack" response lattice. This condition follows directly from the representation of the response lattice as a superposition of the frequency-shifted arrays corresponding to a particular sequence and the requirement that none of the elements of the shifted arrays coincide (if they did, the response lattice would have some lattice points with strengths greater than one). From the earlier discussion of construction of the response lattice from the frequency-shifted arrays $F^{(j)}$ it can be seen that if the distance between each 1 in the F-matrix and all 1's with smaller column numbers is represented as a two-tuple, where the first number is the difference in row number and the second is the difference in column number, a necessary and sufficient condition for a F-matrix to correspond to a "thumbtack" response function is that all of the two-tuples be distinct. If any of the two-tuples are not distinct, there will be a combination of frequency shift and time delay that will cause those lattice points to coincide. By the same reasoning, if two or more lattice points coincide, then two or more of the two-tuples are not distinct.

If the two-tuples are distinct then the other sequences derived by transformation of the F-matrix will also have the "thumbtack" property since the corresponding response lattices are just rotations and/or time-reversals of the original.

In practice, however, it is both easier and more informative to calculate the d_{ij} array, which gives explicitly the position of each point of the response lattice. By exhaustive calculation of the d_{ij} array corresponding to each of the $6! = 720$ sequences of length 6, all 116 distinct possible "thumbtack" sequences may be found. Each of these sequences can then be classified into one of seventeen groups of at most

eight distinct sequences by the methods of the previous section. Table 3.4 is a listing of these "thumbtack" sequences. Because some sequences are transformed into themselves by these operations not all of the groups have the maximum of eight distinct sequences.

3.7 Some Necessary Conditions on the Response Lattice for Realizable Sequences

Although, as in Section 3.5, sequences with particular properties may be found by exhaustive search of all possible sequences of a given length, there is no guarantee that for sufficiently restrictive constraints any sequences which satisfy the constraints can be found. If this is the case it must be concluded that either the constraints are such that no sequences of any length can satisfy these constraints and that the search be terminated or that sequences of greater length which have not been examined might contain a sequence which satisfies the constraints. This trial-and-error approach to waveform design, even though aided by the algorithms developed in Section 3.4, is thus not satisfactory. Thus it is appropriate to consider non-realizable sequences. In this section the d_{ij} array introduced in Section 3.4 will be the central topic. By analyzing this array a number of necessary conditions for a particular response lattice to represent a realizable sequence will be found. Although not sufficient, these conditions are sufficiently restrictive to be useful in the waveform synthesis problem.

Several relatively simple necessary conditions follow directly from the definition of the sequence Γ_n. Since, if a sequence has M elements, there will be a particular frequency shift that will bring γ_1 into the filter to γ_2, and vice-versa there must be a lattice point at $\tau = \pm M - 1$. Because γ_1 and γ_m are the only elements of Γ_n with a spacing in the sequence of $M - 1$, there can only be one lattice point with a delay

$$\tau = M - 1$$

and one point with a delay

$$\tau = (M - 1) .$$

Table 3.4

ALL POSSIBLE SEQUENCES OF LENGTH SIX WITH A "THUMBTACK" RESPONSE LATTICE

Fundamental Sequence	Time reverse	Comple-ment	Time-Reverse Complement	Delay Sequence	Delay Complement	Delay Time Reverse	Delay Time Reverse Complement
231546	645132	546231	132654	312546	465231	645213	132564
512436	634215	265341	143562	235416	542361	614532	163245
413256	652314	364521	125463	243156	534621	651342	126435
351426	624153	426351	153624	351426*	426351*	624153*	153624*
513426	624315	264351	153462	253416	524361	614352	163424
154236	632451	623641	145326	145326*	632451*	523541*	154236*
423516	615324	354261	162453	523146	254631	641325	136452
532416	614235	245361	163542	532416*	245361*	614235*	163542*
452135	631254	325641	146523	435126	342651	621534	156243
524316	613425	253461	164352	524316*	253461*	613425*	164352*
354126	621453	423651	156324	451326	326451	623154	154623
315462	264513	462315	513264	261435	516342	534162	243615
421563	356124	356214	412653	326145	451632	541623	236154
245163	361542	532614	416235	416235*	361542*	532614*	245163*
351264	462153	426513	315624	341625	436152	526143	251634
352164	461253	425613	316524	431625	346152	526134	251643
452613	316254	325164	461523	536124	241653	421635	356142

*Duplicated elsewhere in the group--not a distinct sequence.

However, the element pairs γ_m, γ_2 and γ_{m-1}, γ_1 have a sequence spacing of $M - 2$, which means that there will be two lattice points at

$$\tau = M - 2$$

and two lattice points at

$$\tau = -(M - 2) .$$

By extending this argument to all possible element pairs it can be seen that for a sequence Γ_n with M elements the response lattice will have the following property:

> If $|\tau| = M - j$, there will be exactly j lattice points with this delay (where j is an integer such that $1 \leq j \leq M$).

There is a similar property for the distribution of lattice points as a function of normalized frequency offset ϕ which follows from the above property and the rotation properties of Section 3.5. If $T_n' = \{\tau_1', \tau_2', \ldots, \tau_M'\}$ is the complementary delay sequence corresponding to Γ_n, then the response lattice for T_n' must have the above property. However, the response lattice corresponding to T_n' is a 90° rotation of the response lattice corresponding to Γ_n, and any property of the response lattice corresponding to Γ_n, and any property of the response lattice in the τ direction must then also hold in the ϕ direction. Thus it follows that

> If $|\phi| = M - k$, there will be exactly k lattice points at this normalized frequency offset (where k is an integer such that $1 \leq k \leq M$).

Furthermore, since no response resulting from a frequency shifted subpulse can occur at a delay greater than or equal to the sequence length it follows that

If the normalized time delay τ is such that

$$|\tau| \geq M ,$$

then the response lattice is identically zero.

There is a corresponding property which limits the extent of the response lattice in the ϕ direction:

If the normalized frequency offset ϕ is such that

$$|\phi| \geq M$$

then the response lattice is identically zero.

These properties imply that the response lattice is bounded in the τ and ϕ directions by a square with sides of length $2M$ centered at $\tau = 0$, $\phi = 0$, and that within this square there will be

$$M + 2(M - 1) + 2(M - 2) + \ldots + 2 = M^2 \qquad (3.27)$$

lattice points with M of them located at $\tau = 0$, $\phi = 0$ with the remaining points satisfying the first two properties.

So far, however, the properties which have been found are not particularly useful in actually attempting to synthesize a sequence from a given response lattice.

In order to investigate the synthesis problem more deeply it will be useful to re-introduce the distance array, $[d_{ij}]$, discussed in Section 2.4. Each element of the distance array d_{ij} was defined as the normalized delay n between $\gamma_k = j$ and $\gamma_{k+n} = i + j$ where the frequency offset was assumed positive. However, the delay sequence corresponding to Γ_n,

$$\underline{\tau}_n = \{\tau_1, \tau_2, \ldots, \tau_M\} , \qquad (3.28)$$

is defined by the property that

$$\tau_j = k \quad \text{if and only if} \quad \gamma_k = j \ . \tag{3.29}$$

Then, if $\gamma_k = j$ and $\gamma_{k+n} = i + j$ it follows that

$$\tau_{i+j} - \tau_j = (k + n) - k = n \ . \tag{3.30}$$

Thus, from the definition of d_{ij}, it can be concluded that

$$d_{ij} = \tau_{i+j} - \tau_j \ . \tag{3.31}$$

We need one more property of the d_{ij} array before developing the main result of this section. From the definition of the d_{ij} array it follows that the row sums are given by:

$$S_1 = \sum_{j=1}^{M-1} d_{1j} = \sum_{j=1}^{M-1} [\tau_{1+j} - \tau_j] = \tau_M - \tau_1$$

$$S_2 = \sum_{j=1}^{M-2} d_{2j} = \sum_{j=1}^{M-2} [\tau_{2+j} - \tau_j] = \tau_M + \tau_{M-1} - (\tau_1 + \tau_2)$$

$$S_3 = \sum_{j=1}^{M-3} d_{3j} = \sum_{m=1}^{M-3} [\tau_{3+j} - \tau_j] = \tau_M + \tau_{M-1} + \tau_{M-2} - (\tau_1 + \tau_2 + \tau_3)$$

$$\vdots$$

$$S_k = \sum_{j=1}^{M-k} d_{kj} = \sum_{j=1}^{M-k} [\tau_{k+j} - \tau_j] = \sum_{i=M-k+1}^{M} \tau_i - \sum_{i=1}^{k} \tau_i \tag{3.32}$$

where

$$1 \leq k \leq M - 1 \ .$$

From the above set of equations it can be seen that

$$S_k = S_{M-k} \ . \tag{3.33}$$

Thus, if M is an even integer there will be $M/2$ independent sums; if M is an odd integer there will be $(M - 1)/2$ independent sums.

These row sums have a very important physical significance: S_k is the sum of the time delays of all lattice points with a frequency offset of k. By solving for as many τ_i as possible in terms of the S_k it will be possible to restrict the values which the τ_i may have. Despite the complex appearance of the above set of equations, this solution is easily obtained by recognizing that

$$S_k = S_{k-1} + \tau_{M-k+1} - \tau_k \ . \tag{3.34}$$

Thus all of the delays can be expressed as

$$\tau_{M-k+1} - \tau_k = S_k - S_{k-1} \ , \tag{3.35}$$

where for purposes of consistency S_0 is defined to be identically zero. If M is an even integer the above equation yields a set of $M/2$ independent equations in M unknowns; if M is an odd integer then there will be $M-1/2$ independent equations in $M-1$ unknowns ($\tau_{M+1/2}$ is not defined by these equations if M is odd.)

Since Γ_n is the delay sequence corresponding to T_n, an expression analogous to the above can be written for the row sums, $\hat{S}_k$, of the distance array $\hat{d}_{ij}$ corresponding to T_n as

$$\gamma_{M-k+1} - \gamma_k = \hat{S}_k - \hat{S}_{k-1} \quad . \tag{3.36}$$

Using the property that the response matrix corresponding to T_n is identical to the response matrix of Γ_n except that it has undergone a time reversal followed by a 90° rotation, the row sums S_k can be interpreted as the sum of the frequency offsets of all lattice points of the response lattice corresponding to Γ_n with a normalized delay of k. To solve for Γ_n given a response lattice, these two interlocking sets of equations, one relating the frequency offsets at a particular delay to the γ_i and the other relating to the delays at a particular frequency offset to the τ_i, must be solved subject to the following conditions:

1) $\tau_j = k$ if and only if $\gamma_k = j$; (3.37a)

2) $\tau_i \neq \tau_j$ if $i \neq j$, $i, j = 1, 2, \ldots, M$ (3.37b)

3) $\gamma_i \neq \gamma_j$ if $i \neq j$, $i, j = 1, 2, \ldots, M$ (3.37c)

4) $1 \leq \tau_i \leq M$ (3.37d)

5) $1 \leq \gamma_i \leq M$. (3.37e)

To illustrate how the results of this section can be used to synthesize a sequence from a given response lattice the example below will derive the sequence corresponding to the response lattice in Fig. 3.13.

Example 3.7.1

The first step in this synthesis should be to check the response lattice to guarantee that it has the elementary properties presented at the beginning of this section. By inspecting the response lattice it can be seen that

1. it is skew-symmetric;
2. it is bounded by a square 12 units by 12 units (that $M = 6$ follows from the height of the central response);

3. there are five lattice points with $|\emptyset| = 1$, four with $|\emptyset| = 2$, etc.;

4. there are five lattice points with $|\tau| = 1$, four with $|\tau| = 2$, etc.;

5. $S_1 = S_5 = 5 + 2 - 1 - 2 - 3 = 1$, $S_2 = S_4 = 5 + 2 - 1 - 3 = 2$, and $S_3 = 4 + 2 + 1 = 7$; and,

6. $\hat{S}_1 = \hat{S}_5 = 5 + 3 - 1 - 2 - 4 = 1$, $S_2 = S_4 = 3 + 2 \cdot 1 - 1 = 5$, and $\hat{S}_3 = 4 - 1 - 2 = 1$.

Since the response lattice satisfies the necessary conditions, the synthesis process using the set of simultaneous equations can be continued. Had any of the above tests failed it could be concluded that a sequence corresponding to this particular response lattice does not exist.

Writing the equations for the τ_i and γ_i we obtain:

$$\gamma_6 - \gamma_1 = \hat{S}_1 = 1 , \tag{3.38a}$$

$$\gamma_5 - \gamma_2 = \hat{S}_2 - \hat{S}_1 = 4 , \tag{3.38b}$$

$$\gamma_4 - \gamma_3 = \hat{S}_3 - \hat{S}_2 = -4 , \tag{3.38c}$$

$$\tau_6 - \tau_1 = S_1 = 1 , \tag{3.38d}$$

$$\tau_5 - \tau_2 = S_2 = S_1 = 1, \text{ and} \tag{3.38e}$$

$$\tau_4 - \tau_3 = S_3 - S_2 = 5 . \tag{3.38f}$$

Since

$$\tau_4 - \tau_3 = 5 ,$$

it follows by Eq. (3.37d) that

$$\tau_4 = 6, \quad \tau_3 = 1 .$$

By using the relationship in Eq. (3.37a) we obtain that

$$\gamma_6 = 4, \quad \gamma_1 = 3 \ .$$

Substituting these values for γ_6 and γ_1 into Eq. (3.38a) does not lead to any contradiction, and so we continue. There are two pairs of values for γ_5 and γ_2 which satisfy Eq. (3.38b)

$$\gamma_5 = 6, \quad \gamma_2 = 2$$

or

$$\gamma_5 = 5, \quad \gamma_2 = 1 \ .$$

Using the first pair implies that

$$\tau_6 = 5, \quad \tau_2 = 2 \ .$$

Substituting these values into Eqs. (3.38d) and (3.38e) gives

$$\tau_1 = \tau_6 - S_1 = 4,$$

and

$$\tau_5 = \tau_2 + S_2 - S_1 = 3 \ .$$

We can thus conclude that

$$\gamma_3 = 5, \quad \gamma_4 = 1 \ .$$

Therefore one possible sequence which satisfies the constraints is

$$\Gamma_1 = \{3, 2, 5, 1, 6, 4\} \ .$$

Had we instead chosen

$$\gamma_5 = 5, \quad \gamma_2 = 1 \ ,$$

we would have obtained

$$\tau_5 = 5, \quad \tau_1 = 2 \ .$$

Substituting these values into Eqs. (3.38d) and (3.38e) gives

$$\tau_2 = \tau_5 + S_1 - S_2 = 4,$$

and

$$\tau_6 = \tau_1 + S_1 = 3 \ ,$$

which in turn implies that

$$\gamma_4 = 2, \quad \gamma_3 = 6 \ .$$

Therefore, there is another sequence which satisfies the constraints:

$$\Gamma_2 = \{3, 1, 6, 2, 5, 4\} \ .$$

But, from Table 3.3 and Fig. 3.10 it can be seen that Γ_1 and Γ_2 have the same response lattice and that Γ_1 and Γ_2 are the correct sequences. Since this synthesis examines all γ_i combinations which satisfy the constraints it can also be concluded that the only sequences which yield the response lattice of Fig. 3.13 are Γ_1 and Γ_2.

This chapter has considered the interaction of the response lattice and the sequence which generates it. The next chapter will discuss error rates which can be realized with these sequences in channels subject to rapid fluctuation and the generation of sets of sequences to minimize these error rates.

Chapter IV

MODULATION TECHNIQUES AND ERROR RATES FOR FADING CHANNELS

4.1 Introduction

In this chapter we will analyze techniques suitable for transmitting digital data, with particular emphasis on techniques suitable for rapidly varying channels. In this analysis we will compare error rates achieved by standard modulation techniques such as frequency shift keying (FSK), phase shift keying (PSK), and differentially coherent phase shift keying (DPSK) under conditions of "slow" and "fast" fading and the improvement in error rate provided by diversity. These techniques will then be compared with the incoherently detected coded frequency shift keyed sequences introduced in Chapter III, and the problem of optimizing these sequences for particular channel characteristics will be discussed in detail. We will find that the error rates in rapidly fading channels are sequence dependent and that the response lattice concept introduced in Chapter III is convenient for representing these dependencies.

4.2 Error Rates for Slow Fading--No Diversity

The effect of a multiplicative time-varying channel on a transmitted signal $u(t)$ can be compactly represented by complex waveform notation as

$$Y(t) = Z(t)\, U(t) \ , \tag{4.1}$$

where $Y(t)$, $Z(t)$, and $U(t)$ are, respectively, the complex envelopes of the received, channel corrupted waveform, the channel process, and the transmitted signal

$$u(t) = \mathrm{Re}\{U(t) \exp [j2\pi ft]\} \ . \tag{4.2}$$

The statistics of $Y(t)$ are the fading statistics; the rate of change of $Y(t)$ compared to the length of time coherent summation (for example, in

a matched filter) of the received waveform is being attempted and the desired probability of error will determine the fading rate for which the fading can be considered "slow."

We will defer until later a more precise definition of slow fading; however, if the fading is sufficiently slow, then the probability of error is obtained by averaging the probability of error conditioned on the received signal strength over the range of signal strengths given by the channel statistics. If we also assume that the channel statistics are adequately represented by a Rayleigh distribution with mean signal-to-noise ratio γ_0 then the average probability of error of several common modulation and detection schemes is easily calculated. It can be shown [Schwartz, Bennett, and Stein, 1966] that the conditional probability of error for common FSK and PSK systems is

$$P_{e,\gamma}^{(1)} = \frac{1}{2}\exp(-a\gamma), \quad \text{where} \quad \begin{cases} a = \frac{1}{2}, & \text{noncoherent FSK} \\ a = 1, & \text{DPSK} \end{cases} \tag{4.3a}$$

or

$$P_{e,\gamma}^{(2)} = \frac{1}{2}\operatorname{erfc}(\sqrt{a\gamma}), \quad \text{where} \quad \begin{cases} a = \frac{1}{2}, & \text{coherent FSK} \\ a = 1, & \text{coherent PSK} \end{cases} \tag{4.3b}$$

Averaging these conditional probabilities of error over the assumed Rayleigh density gives

$$P_e^{(1)} = \int_0^\infty \left[\frac{1}{2}\exp(-a\gamma)\right]\left[\frac{1}{\gamma_0}\exp\left(\frac{-\gamma}{\gamma_0}\right)\right] d\gamma = \frac{1}{2}\left[\frac{1}{1 + a\gamma_0}\right] \tag{4.4a}$$

or

$$P_e^{(2)} = \int_0^\infty \left[\frac{1}{2}\operatorname{erfc}(\sqrt{a\gamma})\right]\left[\frac{1}{\gamma_0}\exp\left(\frac{-\gamma}{\gamma_0}\right)\right] = \frac{1}{2}\left[1 - \frac{1}{\sqrt{1 + 1/a\gamma_0}}\right] \tag{4.4b}$$

Thus we obtain the following average error probabilities:

$$P_e = \frac{1}{2 + \gamma_0} \quad , \quad \text{noncoherent FSK} \tag{4.5a}$$

$$P_e = \frac{1}{2}\left[1 - \frac{1}{\sqrt{1 + 2/\gamma_0}}\right] , \quad \text{coherent FSK} \tag{4.5b}$$

$$P_e = \frac{1}{2}\left[1 - \frac{1}{\sqrt{1 + 1/\gamma_0}}\right] , \quad \text{coherent PSK} \tag{4.5c}$$

$$P_e = \frac{1}{2 + 2\gamma_0} \quad , \quad \text{DPSK} . \tag{4.5d}$$

These error probabilities are plotted in Fig. 4.1. For $\gamma_0 > 10$ dB the curves are closely approximated by straight lines of slope minus one. Therefore, the probability of error for slow Rayleigh fading channels is proportional to $1/\gamma_0$, in contrast to the exponential decrease with γ_0 observed in nonfading channels.

It is apparent from these curves that a small probability of error requires quite large signal-to-noise ratios. To avoid this large SNR requirement some type of diversity is usually employed for fading channels. Several of the more common types of diversity and the effect of diversity upon error rates will be discussed in the next section.

4.3 Diversity Techniques

We found in Section 4.2 that a low error rate in a Rayleigh fading channel requires a considerably increased signal-to-noise ratio over that required by a nonfading channel. Intuitatively, the increased SNR requirement reflects the high error rates which occur during fading nulls. To reduce the percentage of time that the channel has faded to a certain received signal level and thus reduce the error rate, the mean signal-to-noise ratio must be increased. Alternatively, if a second channel whose

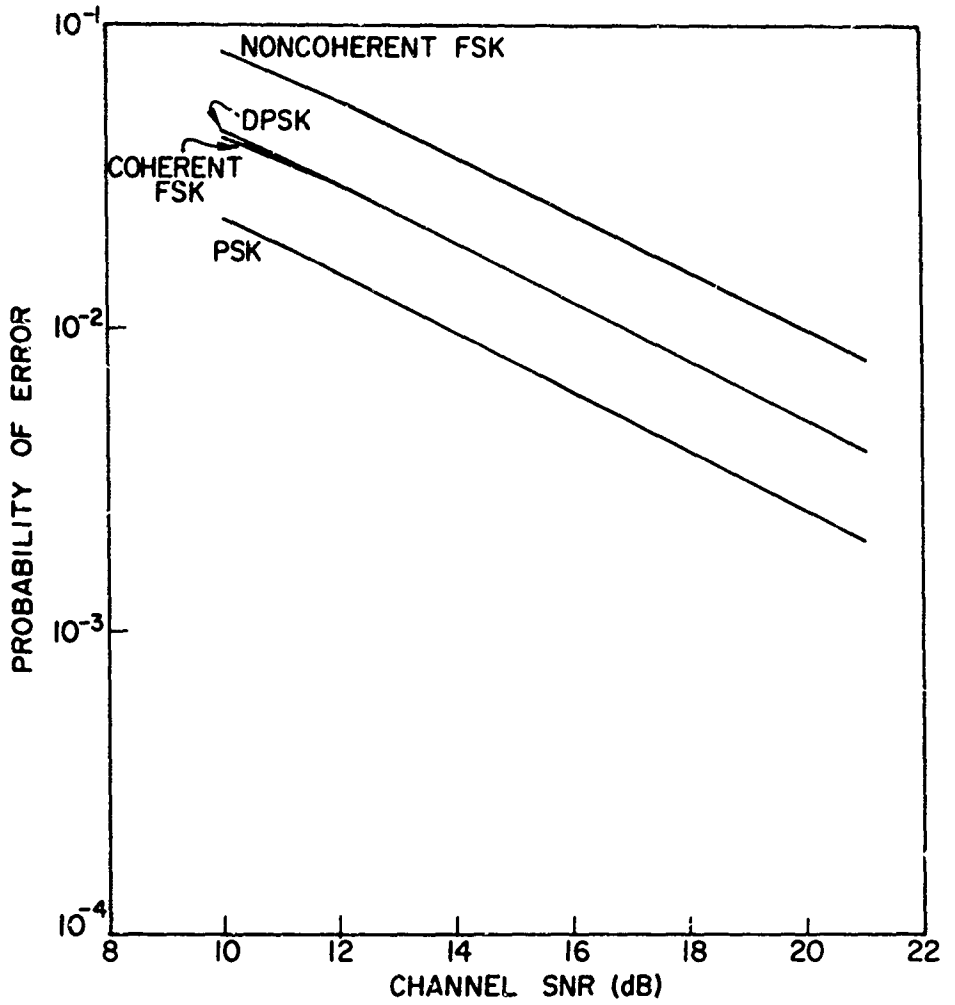


Fig. 4.1. PROBABILITY OF ERROR FOR RAYLEIGH CHANNEL WITH SLOW FLUCTUATIONS.

fades are independent of the fades in the first can be found, then the same message can be transmitted over both channels. Since the channels are independent, the probability that both signals simultaneously fade below a certain level is reduced. This same argument can be extended to any number of diversity channels. However, beyond a certain number of diversity channels there is little to be gained by adding more. In fact, for certain types of energy-sharing diversity (to be explained later) there is an optimum number of diversity channels; adding more channels than is optimum will increase the error rate.

There are a large number of techniques which can be employed to produce channels with approximately independent fading statistics. The choice of which technique or combination of techniques to use will depend upon the physical characteristics of the channel, the desired error rate, and the allowable equipment complexity. However, most diversity techniques can be classified into three basic categories. The first of these basic classes is space diversity. This technique uses two or more receiving antennas spaced a number of wavelengths apart. It was found experimentally that the fading on antennas with a sufficiently large spacing was nearly decorrelated, apparently because the signals arriving at the two antennas have traveled over different paths. As long as the received signals remain decorrelated, a number of receiving antennas may be used. A variation on this technique is to provide both multiple receiving and multiple transmitting antennas. A second, commonly used method for obtaining diversity, called frequency diversity, is to transmit at several widely spaced frequencies. As long as the frequency spacing is sufficiently large, the signal received on each of these frequencies will fade independently. A third technique for providing diversity is to repeat a message several times, on the assumption that if the spacing between messages is sufficiently long, then the channel fading level will be independent from message to message. This is a simple example of time-diversity.

Many of the more exotic diversity techniques proposed or in use are combinations or variations of these three elementary methods. For example, the Rake system (Price and Green [1958]) transmits a signal with sufficient

bandwidth to resolve the difference in arrival time at the receiver of signals which have traversed paths of different length. These components, since they have traveled over different paths, can be expected to fade independently. Thus the Rake system uses a form of space diversity where the diversity channels are separated on the basis of arrival time rather than spatial location.

In HF communications the bandwidth over which the fading is correlated may be less than the information bandwidth; thus the two sidebands of a double-sideband AM signal are sufficiently widely spaced to provide frequency diversity. Another similar variation is to amplitude modulate a carrier with a number of tone bursts. If the frequency separation of these tone bursts is comparable to the fading correlation bandwidth, then the fading of each modulating tone will be approximately uncorrelated, providing frequency diversity. By sending the tone bursts sequentially rather than simultaneously and recombining in the receiver it is possible to realize time diversity as well as frequency diversity.

Coding, particularly coding for burst errors, is a nontrivial form of time-diversity. By knowing or measuring the fading statistics of the channel, it is possible to distribute the information content of the message so that with high probability enough of the encoded bits can be correctly received to permit correct decoding of the message.

In any particular system the actual technique used to obtain diversity has no effect on error rates provided the different techniques yield the same mean signal-to-noise ratio. Thus, error rates as a function of the number of independent diversity branches can be calculated without having to consider what type of diversity is actually being employed. In the next section we will derive error rates for fluctuating channels as a function of the number of diversity branches.

4.4 Error Rates for Arbitrary Fading Rate and Order of Diversity-Binary Alphabet

If the rate at which the channel process fluctuates approaches the data bandwidth the analysis of Section 4.2 is no longer valid. By assuming that the received time-varying signal, including additive thermal

noise, is representable as a complex, nonstationary gaussian process, Bello and Nelin [1962] derived an expression relating the probability of error for a binary communication system to the fading rate and order of diversity. They assume that the decision as to which of the two possible messages was actually sent is made by a threshold decision on a random variable q, where, in general,

$$q = \sum_{k=1}^{L} \left[a|W_k|^2 + b|X_k|^2 + cW_k^* X_k + c^* W_k X_k^* \right] \tag{4.6}$$

and W_k and X_k are the complex outputs of filters in the kth diversity branch tuned to the mark and space symbols, respectively.

By choosing a, b, and c appropriately the quadratic form for q given above can represent coherent, differentially coherent, and incoherent detection, as well as either pre-detection or post-detection combining. However, recalling the discussion in Chapter II relating to the difficulty of deriving a phase reference in a rapidly fluctuating channel, we will confine the discussion in this section to incoherent detection and post-detection combining of the diversity channels. For incoherent detection $a = 1$, $b = -1$, and $c = 0$. With these restrictions the decision variable q becomes

$$q = \sum_{k=1}^{L} \left[|W_k|^2 - |X_k|^2 \right] .$$

A block diagram for this incoherent matched filter receiver is shown in Fig. 4.2.

Errors may be made in two ways: either q may be positive given that a "space" symbol was sent or q may be negative given that a "mark" symbol was sent. These two probabilities will be represented by p_0 and p_1, respectively. Bello and Nelin show that the probability of these two types of errors is

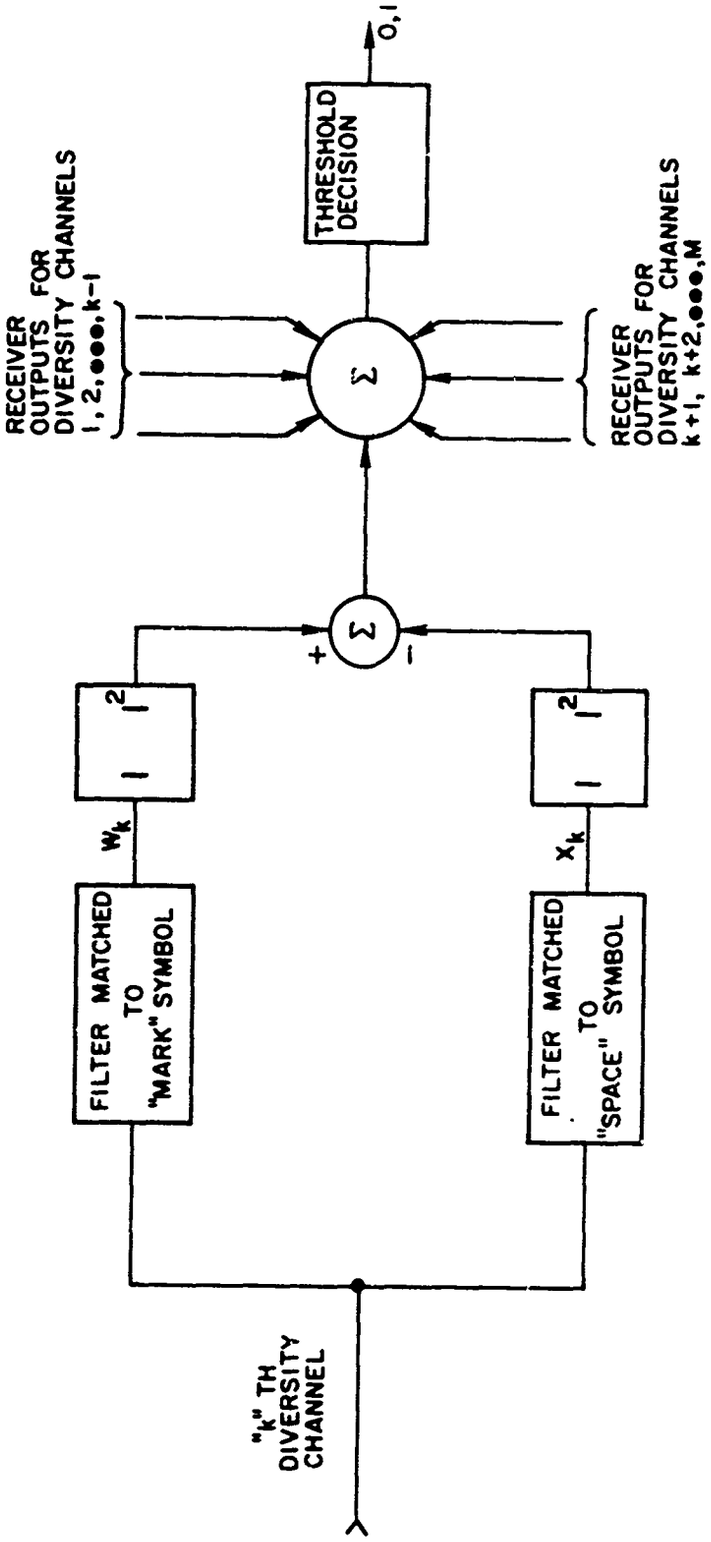


Fig. 4.2. INCOHERENT MATCHED FILTER DIVERSITY RECEIVER.

$$p_0 = p_r(q > 0 | \text{space}) = \left[\frac{1 + \beta_0}{2 + \beta_0}\right]^L \sum_{m=0}^{L-1} \left[\binom{L-1+m}{L-1} \frac{1}{(2 + \beta_0)^m}\right] \quad (4.7)$$

and

$$p_1 = p_r(q < 0 | \text{mark}) = \frac{1}{(2 + \beta_1)^L} \sum_{m=0}^{L-i} \left[\binom{L-1+m}{L-1}\left(\frac{1 + \beta_1}{2 + \beta_1}\right)^m\right]. \quad (4.8)$$

For a slowly fluctuating channel with signals of equal energy $p_1 = p_0$; however, rapidly fluctuating channels may have $p_1 \neq p_0$ unless the fading spectrum is symmetrical and the mark and space waveforms have identical envelopes. In these expressions β_i represents a generalized diversity branch signal-to-noise ratio given by

$$\beta_i = \frac{\left(2\, m_{11}^i - m_{00}^i\right)}{\sqrt{\left(m_{11}^i - m_{00}^i\right)^2 + 4\left(m_{11}^i m_{00}^i - |m_{10}^i|^2\right)} - \left(m_{11}^i - m_{00}^i\right)}, \quad (4.9)$$

where $i = 1$ for "mark" transmitted, $i = 0$ for "space" transmitted,

$$m_{11}^i = \overline{|W_k|^2}, \quad m_{10}^i = \overline{W_k^* X_k}, \quad m_{00}^i = \overline{|X_k|^2} \quad (4.10)$$

and all other moments such as $\overline{|W_k W_j^*|}$, $\overline{|X_k X_j^*|}$, and $\overline{|W_k^* X_j|}$ are assumed equal to zero. These restrictions mean that each diversity branch is independent, since the outputs of the filters are assumed gaussian. The general form of the expression for the moments is

$$m_{rs}^i = \int_{-T}^{T} R(\tau)\, \lambda_{rs}^i(\tau) + 4N_0 E_{rs}, \quad (4.11)$$

where N_0 is the additive noise density, $S_i(t)$ is the transmitted symbol,

$$R(\tau) = \overline{Z_k^*(t_1)\, Z_k(t_1 + \tau)} \,, \tag{4.12}$$

$$\lambda_{rs}^{i}(\tau) = \int_0^T S_i^*(t)\, S_r(t)\, S_i(t + \tau)\, S_s^*(t + \tau)\, dt \,, \tag{4.13}$$

and

$$E_{rs} = \int_0^T S_s^*(t)\, S_r(t)\, dt \,. \tag{4.14}$$

Thus the moments depend on the channel fading process $Z_k(t)$, the form and energy of the "mark" and "space" symbols, the additive noise power, and which of the two possible symbols S_0 or S_1 (corresponding to $i=0$ or $i=1$) is being transmitted.

Inspection of the expressions for p_0 and p_1 shows that $p_0 = p_1$ if

$$\beta_1 = -\frac{\beta_0}{1 + \beta_0} \,. \tag{4.15}$$

Thus p_0 can be calculated by first computing β_0 and either substituting directly into the expression for p_0 or by setting

$$\beta_1' = -\frac{\beta_0}{1 + \beta_0} \tag{4.16}$$

and then this quantity into the expression for p_1.

Since the expressions for probability of error are functions of both the transmitter waveform as well as the channel fading statistics it will

be necessary to assume a specific form for these quantities. Because the sequences discussed in Chapter III are a more general form of binary FSK it will be useful to analyze the binary FSK system first. For a binary FSK system the "mark" and "space" waveforms can be represented as:

$$S_1(t) = \sqrt{2E/T}\ \exp\left[-j\left(\frac{n\pi}{T}\right) t\right] ; \qquad (4.17a)$$

$$S_2(t) = \sqrt{2E/T}\ \exp\left[j\left(\frac{n\pi}{T}\right) t\right] , \qquad (4.17b)$$

where E is the waveform energy, T is the waveform duration, and n is the frequency spacing in multiples of $1/T$.

Bello and Nelin analyze two possible forms for $R(\tau)$, the channel fading autocorrelation function: gaussian and exponential. For the low data rate systems (including long pulse radar) considered in Chapter II the exponential fading autocorrelation function should be a reasonable model since one of the principal sources of system instability is the instability of oscillators and random changes in the phase path length. The processes which cause these instabilities can be modeled as a Weiner process (neglecting periodic fluctuations such as might be caused by oscillator power supply ripple), which has an autocorrelation of the form

$$R(\tau) = \exp[-\alpha|\tau|] . \qquad (4.18)$$

Thus it seems reasonable to model to channel fading by a gaussian process with an autocorrelation

$$R(\tau) = e^{\tau 2\pi B|\tau|} , \qquad (4.19)$$

where B is the 3 dB bandwidth of the "single-pole" spectrum corresponding to this autocorrelation function.

Calculating the m_{rs}^{0} for an incoherent FSK system and this particular $R(\tau)$ gives

$$m_{00}^{0} = 8E^2 \left[\frac{1}{2\pi b} + \frac{e^{-2\pi b} - 1}{(2\pi b)^2} + \frac{1}{2\rho} \right] , \tag{4.20a}$$

$$m_{10}^{0} = 8E^2 \left[\frac{e^{-2\pi b} - 1}{(2\pi)^2 (b^2 + n^2)} \right] \tag{4.20b}$$

$$m_{11}^{0} = 8E^2 \left[\frac{1}{(2\pi)^2 (b^2 + n^2)} \left\{ 2\pi b - \frac{b^2 - n^2}{b^2 + n^2} (1 - e^{-2\pi b}) \right\} + \frac{1}{2\rho} \right] \tag{4.20c}$$

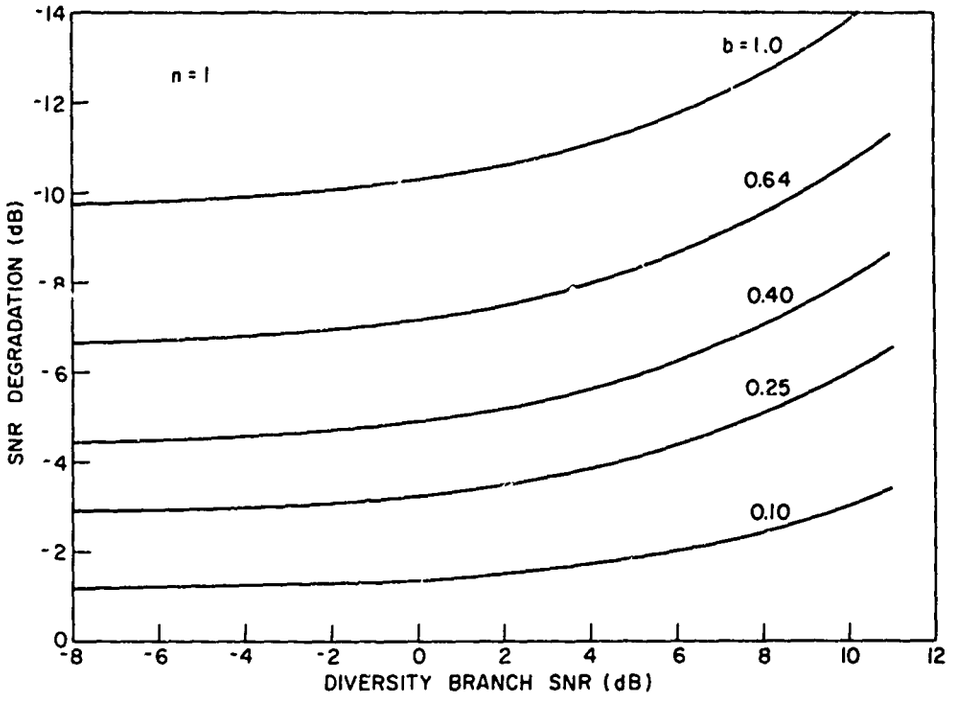
where ρ is the diversity channel SNR and

$$b = BT .$$

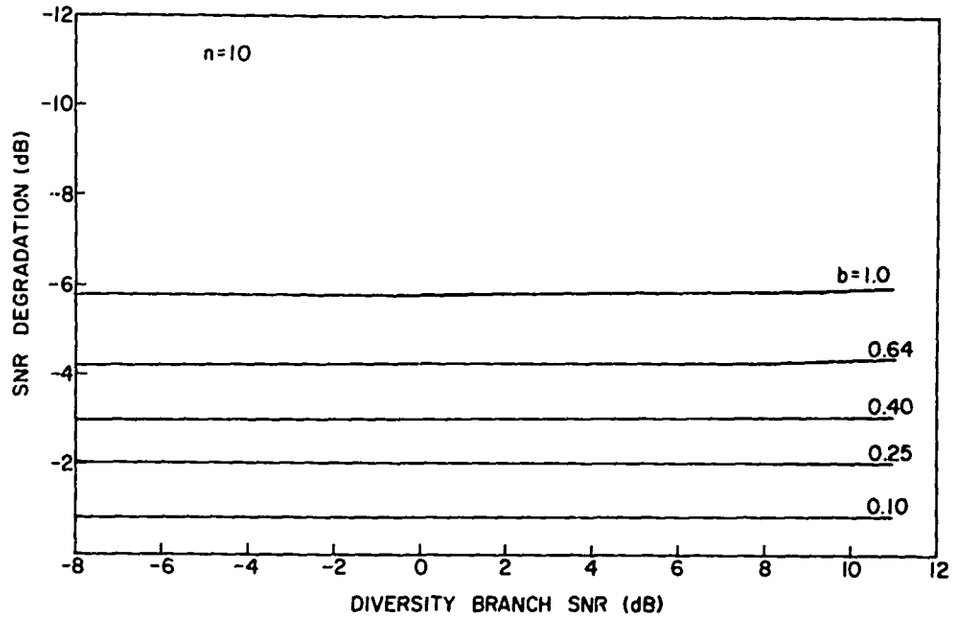
For particular values of b, n, and ρ these expressions permit the calculation of the moments. These moments are substituted into Eq. (4.9) to obtain β_0, which allows p_0 to be calculated from Eq. (4.16) and (4.8). It is also possible to calculate p_0 directly from Eq. (4.7).

Since the quantity β_1' can be interpreted as an equivalent signal-to-noise ratio which includes the effect of rapid fading on the channel signal-to-noise ratio under slow, Rayleigh fading, by taking the ratio of β_1' to the slow fading SNR it is possible to plot the effective degradation as shown in Fig. 4.3(a) and Fig. 4.3(b). It can be seen from these curves that the degradation is an increasing function of diversity branch SNR and normalized bandwidth of the channel fading process and a decreasing function of the channel spacing. For sufficiently rapid fading the effective diversity branch SNR can decrease enough to cause a system to be unusable because of the high error rate.

There are two sources for this degradation: first, the channel spreads some of the signal energy outside the bandpass of the matched filter, and, second, energy from the "mark" waveform is spread into the filter for the "space" waveform, and vice versa. The first effect could be



(a)



(b)

Fig. 4.3. SNR DEGRADATION FOR EXPONENTIAL FADING AUTOCORRELATION (n = 1, 10.)

overcome simply by increasing the received signal energy; however, increasing the signal energy also increases the amount of energy spread into the filter for the other symbol, preventing, beyond a certain point, any further improvement in error performance.

Using the SNR degradation it is possible to plot the binary error rate as a function of signal-to-noise ratio, number of diversity channels, and the "spread factor" b under the assumption that the total energy/symbol is constant and that the symbol energy is equally divided among the diversity channels. Therefore, because the energy is divided in this manner an increase in the number of diversity channels results in a proportionate decrease in the symbol energy received in each diversity channel. Without this assumption it would be possible to increase the energy/symbol without bound by increasing the number of diversity channels. This restriction is applicable to systems using time and frequency diversity as well as to systems deriving angle diversity from rapidly scanned antennas or space diversity from multiple transmitter antennas.

Figures 4.4(a)-4.4(c) show the probability of error under the assumption of constant energy/symbol for a channel having an exponential fading correlation function with the normalized fading bandwidth b as a parameter. These curves have several interesting properties. For a particular probability of error there is an optimum number of diversity channels which achieve this probability of error with minimum total energy, and the minimum occurs for a diversity channel SNR of about 5 dB. This characteristic of energy-sharing diversity with post-detection combining was first noted by Pierce [Pierce, 1958] for the case of channels with slow fading. In addition, inspection of the curves presented here shows that the optimum number of diversity channels for a particular total SNR is insensitive to the fading bandwidth. As b increases, curves representing low orders of diversity asymptotically approach lines of constant probability of error. Thus for systems with low order diversity it may be impossible to obtain a desired probability of error regardless of how much transmitter power is available. Actually, all of the curves show this saturation effect, but for high order diversity it occurs for error rates that are much lower than usually required.

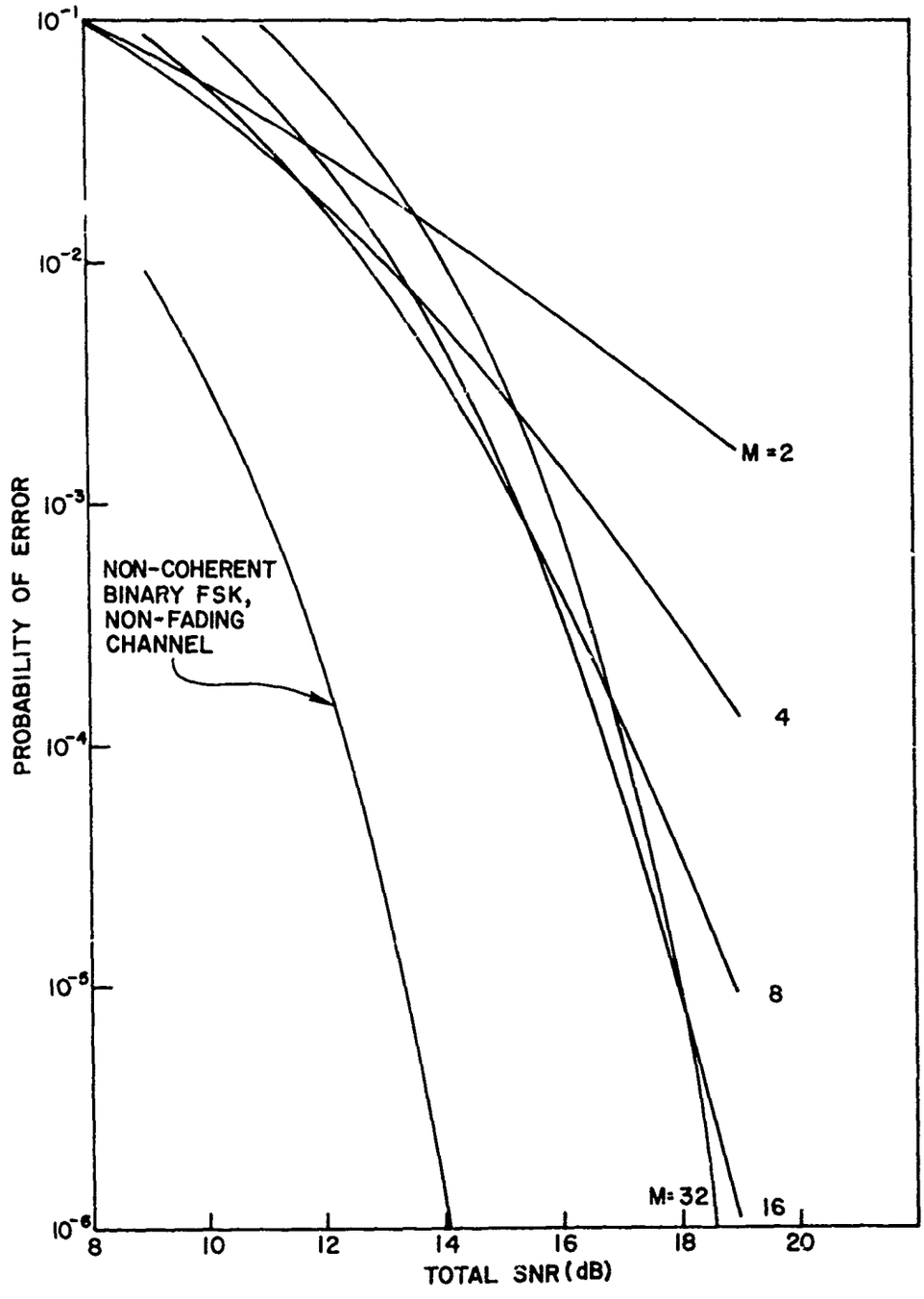
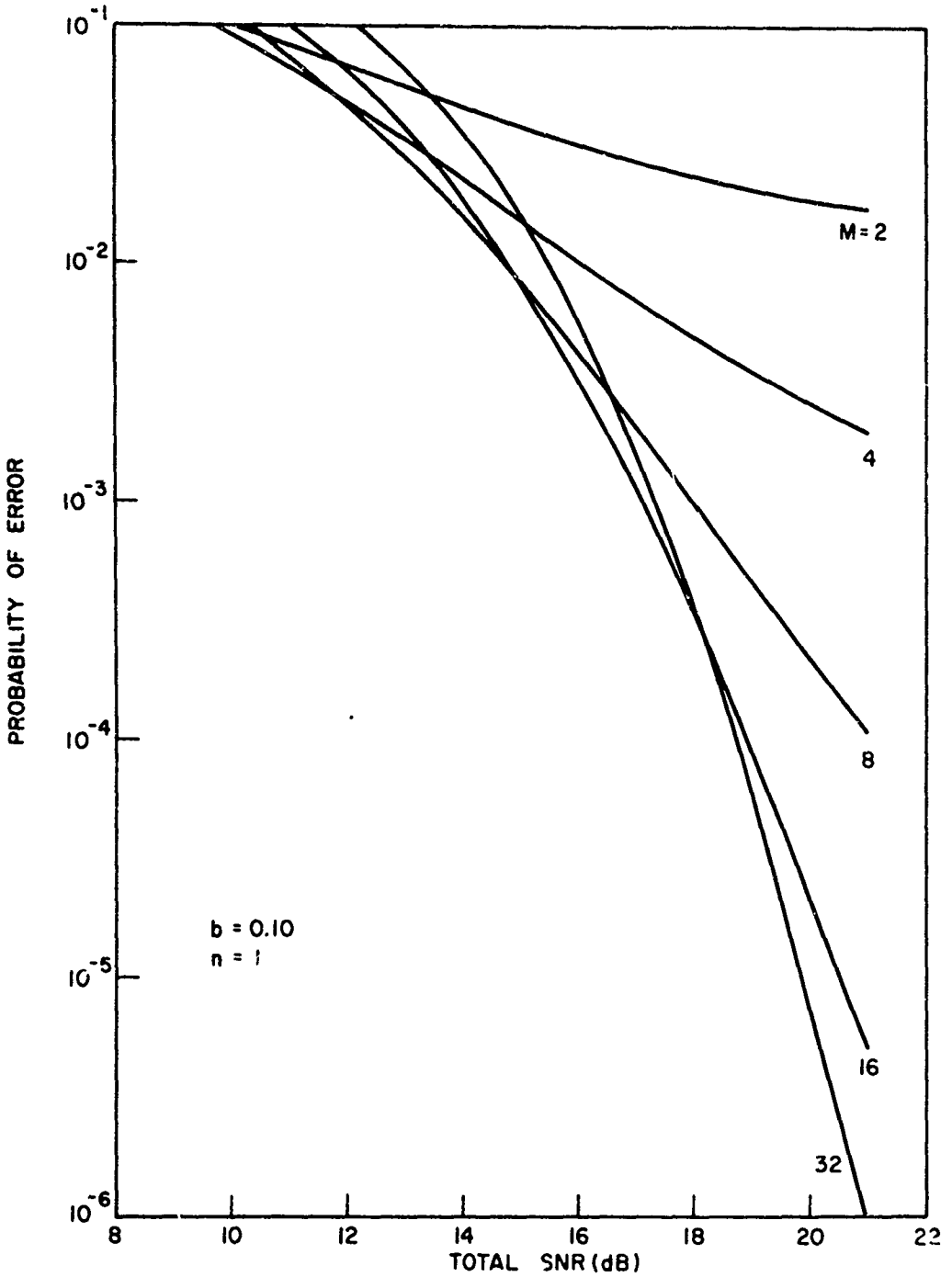


Fig. 4.4a. PROBABILITY OF ERROR VS TOTAL SNR FOR RAYLEIGH CHANNEL WITH SLOW FADING M IS THE NUMBER OF DIVERSITY CHANNELS.



Fib. 4.4b. PROBABILITY OF ERROR VS TOTAL SNR FOR RAYLEIGH CHANNEL WITH EXPONENTIAL FADING AUTOCORRELATION (b = 0.10.)

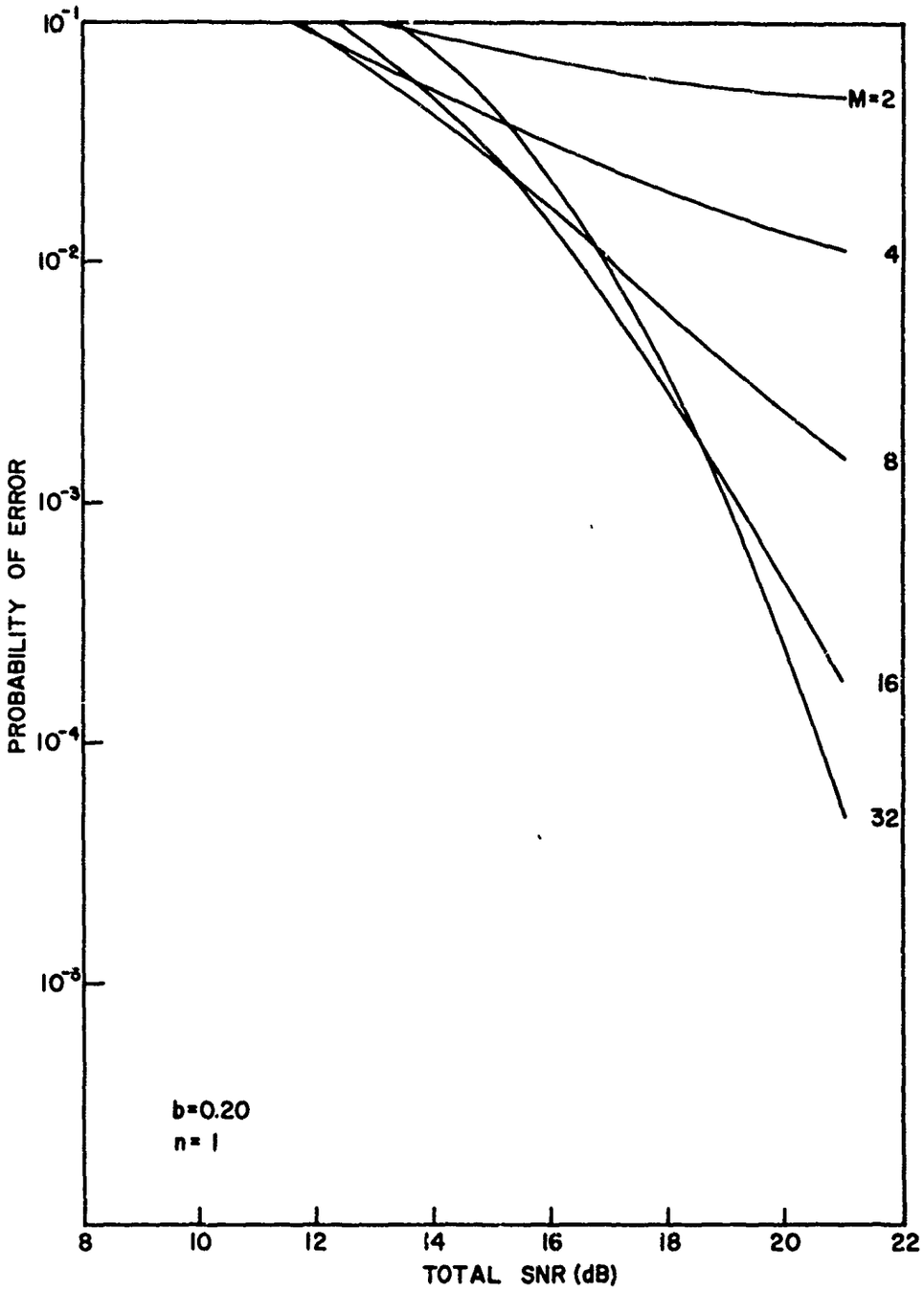


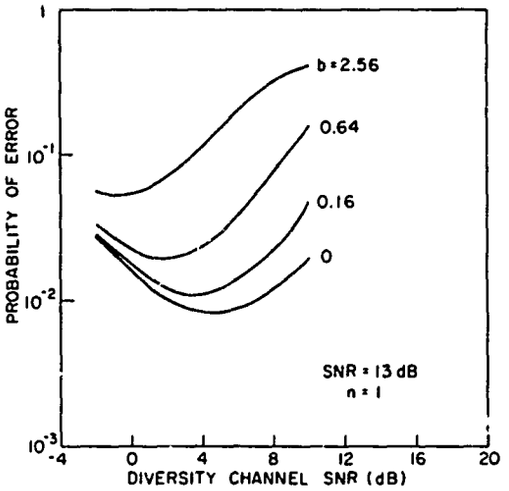
Fig. 4.4c. PROBABILITY OF ERROR VS TOTAL SNR FOR RAYLEIGH CHANNEL WITH EXPONENTIAL FADING AUTOCORRELATION (b = 0.20.)

However, if in contrast with the above discussion the normalized fading bandwidth is a function of the number of diversity channels, then a diversity channel SNR of approximately 5 dB is not necessarily optimum. The next section will discuss the effect of this modification and how it applies to the coded frequency sequences discussed in Chapter III.

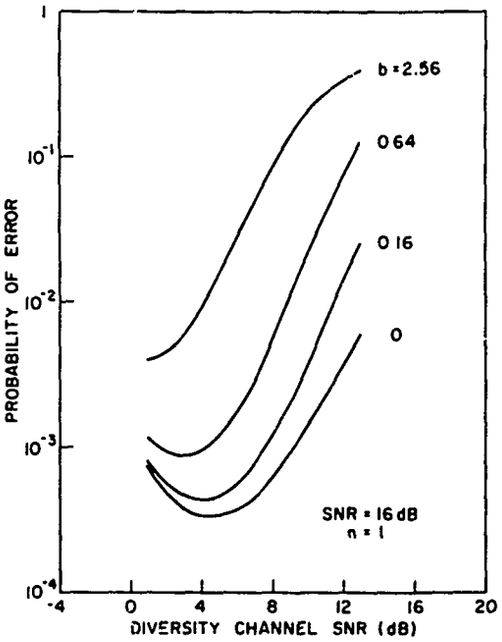
4.5 Optimization of Diversity Channel Power and Coded Frequency Sequences

In deriving the results of Section 4.4 it was assumed that the normalized fading bandwidth in each diversity channel was a constant independent of the number of diversity channels. This assumption would apply to a system where the available power was equally divided among the diversity channels but the duration of transmission remair d constant. If, instead, the total transmitter power is allotted to each diversity channel for a fraction 1/M of the symbol duration a much lower probability of error can be attained for the same total energy in a rapidly fading channel. This benefit obtains because decreasing the duration of the signals in each diversity channel decreases the effective channel fluctuation bandwidth. Thus increasing the number of diversity channels increases the number of independently fluctuating signal paths and decreases the effective fading bandwidth. An example of a system using this second alternative would be a system with M possible transmitter frequencies which transmits at a frequency f_1 for T/M sec, then steps to f_2 for T/M sec, then to f_3 for T/M sec, and continues until all M frequencies have been used. Such a system would use the time and frequency spacing of the diversity channel signals to obtain independently fading paths. The benefit of shortening the diversity channel signals is illustrated in Fig. 4.5(a),(b),(c) which give the probability of error for a rapidly fading Rayleigh channel as a function of BT and diversity channel SNR (which is the total SNR divided by the number of diversity channels). The reason these curves decrease so rapidly as the number of diversity channels increases is that the effective fluctuation bandwidth becomes

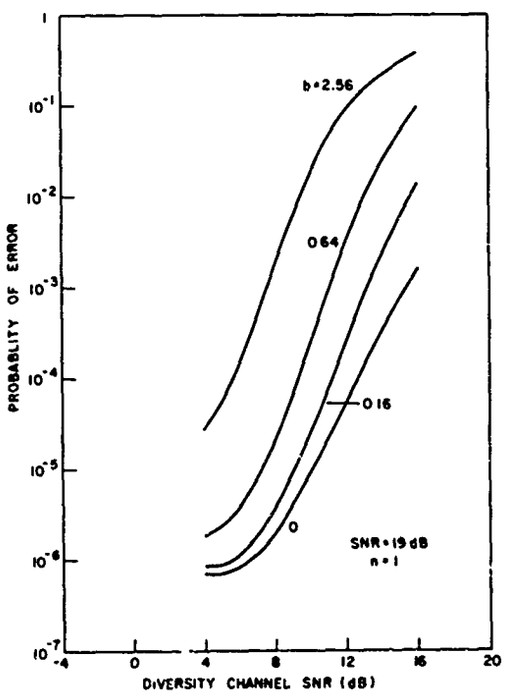
$$b_{eff} = BT/M \ . \tag{4.21}$$



(a)



(b)



(c)

Fig. 4.5. PROBABILITY OF ERROR VS DIVERSITY CHANNEL SNR FOR ENERGY SHARING DIVERSITY AND DIVERSITY CHANNEL SIGNAL DURATION PROPORTIONAL TO 1/M.

However, increasing the number of diversity channels beyond the number corresponding to the minima of the curves causes the square-law detection loss to increase faster than the gain resulting from the increase in diversity and the decrease in fluctuation bandwidth, which causes the error rate to increase again. Fortunately, the minima are sufficiently broad that system performance is not critically dependent on the actual diversity channel SNR.

Inspection of the curves shows that as b increases the minimum of the curve corresponding to this b shifts to a lower value of diversity channel SNR. Thus for a slowly varying channel the optimum diversity channel SNR is about 5 dB, but for $b = 0.64$ the optimum SNR is about 3 dB and for $b = 2.56$ the optimum SNR is only 0 dB. Examination of the curves shows that the channel SNR at which these minima occur depends primarily on b and is nearly independent of the total SNR.

From the viewpoint of minimizing system sensitivity to changes in the fluctuation bandwidth, the system should be designed to operate with a diversity channel SNR of about 3 dB. This choice of SNR permits the fluctuation bandwidth to change over a wide range without greatly affecting the probability of error.

These conclusions about optimum operating points also apply to radar detection of rapidly fluctuating targets; for maximum detectability consistent with minimizing sensitivity to fluctuations with unknown rates the pulse length and pulse duration should be adjusted to yield a single-pulse SNR of approximately 3 dB.

These results are directly applicable to calculation of error rates for systems using the CFSK sequences discussed in Chapter III. As long as the fluctuations of each of the M subpulses are approximately independent as a result of the frequency spacing, time spacing, or both, then the incoherent summation of the outputs of the filters matched to each M frequency shifted subpulses gives, in effect, the equivalent of M diversity channels. If the fluctuations of the subpulses are not independent then the dependencies result in an effective lower order of diversity; however, there must be a high correlation among the diversity channel fading before the diversity improvement is seriously affected.

Turin [1962] calculated the effect of diversity channel correlation for channels with "slow" fading under the assumption that the fading correlation between diversity channels was an exponential function of the channel spacing. This model is consistent with the assumption of an exponential fading autocorrelation function. He showed that even for high correlations considerable improvement can still be obtained; for example, a correlation coefficient $|r|$ of 0.70 results in a 1.8 dB loss of SNR for $M = 10$ and $P_e = 10^{-4}$. For the assumed $R(\tau)$ the normalized fading bandwidth corresponding to this correlation coefficient is

$$b = BT = \frac{1}{2\pi} \ell n \left(\frac{1}{|r|}\right) = 0.057 \qquad (4.22)$$

Thus a small fluctuation bandwidth is sufficient to decorrelate the subpulses enough that nearly the full diversity improvement is obtained. In addition, if the system is operating with approximately the optimum number of diversity channels most of the reduction in error rate is the result of the reduction in b rather than the large number of diversity channels. Therefore, a change in channel statistics causing the subpulse-to-subpulse correlation to increase, decreasing the effective number of diversity channels, will be offset by the decrease in fading bandwidth. Because of this tradeoff the actual effective SNR degradation will be less than that given by Turin's analysis for channels with "slow" fading, and the error rates for CFSK sequences are not critically dependent on the assumption that the subpulses fade independently.

When calculating the probability of error, however, the change in frequency spacing between the "mark" and "space" sequences must be included. For example if we choose the sequence

$$\Gamma_1 = \{1,2,3,4,5,6\}$$

to be the "mark" sequence and

$$\Gamma_2 = \{6,5,4,3,2,1\}$$

to be the "space" sequence then the normalized frequency spacing between Γ_1 and Γ_2 is not constant. Letting n_i be the frequency spacing between the ith subpulse of Γ_1 and Γ_2 gives

$$n_1 = -5 , \quad n_2 = -3 , \quad n_3 = -1 ,$$

$$n_4 = 1 , \quad n_5 = 3 , \quad n_6 = 5 .$$

Thus our assumption of constant frequency spacing between the "mark" and "space" diversity channels is violated. However, a normalized channel spacing of 1 represents the worst case. Since all of the calculations assumed that n=1, the error rates which result are upper bounds on the error rates for particular choices of sequences for the "mark" and "space" symbols. The curve for b=0 is a lower bound. For most purposes these upper bounds will be adequate since the number of diversity channels required for minimum probability of error will result in a low normalized fading bandwidth. By referring to Fig. 5.2(a) and (b) it can be seen that for $b < 0.1$ and a diversity channel SNR of 5 dB or smaller there is less than a 1 dB difference in channel SNR degradation between $n = 1$ and $n = 10$ (which represents essentially infinite symbol spacing). Thus at the error minima the probability of error calculated for n=1 will be a good approximation to the actual, sequence dependent, error probabilities.

A third assumption made in deriving these error probabilities is that the "mark" and "space" symbols are orthogonal. For the sequences considered in Chapter III with frequency spacing constrained to multiples of 1/T the orthogonality assumption is satisfied by any pair of sequences for which

$$n_i \neq 0 , \qquad i = 1,2,\ldots,M .$$

However, this condition for orthogonality is only valid as long as there are no mean frequency shifts. If there are uncompensated mean frequency shifts then sequences which are orthogonal under zero frequency

shift can become highly correlated. The next section will discuss this problem.

4.6 Selection of Orthogonal Sequences

As long as the receiver is exactly tuned to the transmitted sequence the condition discussed at the end of Section 4.5 is sufficient to guarantee that the sequences are orthogonal, thereby minimizing the crosstalk between the "mark" and "space" receivers. However, if there is uncompensated mean doppler shift or oscillator drift that represents a substantial fraction of the subpulse bandwidth M/T then there may be considerable crosstalk even in the absence of rapid fading.

The amount of signal energy which is received in the "space" receiver when a "mark" is transmitted and frequency shifted by an amount

$$\phi = \frac{\Delta f}{M/T} \tag{4.23}$$

is given by the cross-ambiguity function of the "mark" and "space" symbols. To simplify the analysis it will be convenient to define a cross-response lattice analogous to the response lattice representation for the ambiguity function. Thus two sequences Γ_1 and Γ_2 for which the cross-response lattice is desired must first be transformed into their corresponding delay sequences T_1 and T_2. Then the amount of normalized frequency shift before a particular subpulse of Γ_1 aliases as a subpulse of Γ_2 is given by

$$\phi_{ij} = \tau_{i,1} - \tau_{j,2} , \tag{4.24}$$

where $i - j$ is the normalized time delay at which the aliased subpulse is received.

For

$$\Gamma_1 = \{1,2,3,4,5,6\}$$

and

$$\Gamma_2 = \{6,5,4,3,2,1\}$$

the $\{\phi_{ij}\}$ array is easily calculated to be

$$\{\phi_{ij}\}_{\Gamma_1,\Gamma_2} = \begin{Bmatrix} -5 & -4 & -3 & -2 & -1 & 0 \\ -4 & -3 & -2 & -1 & 0 & 1 \\ -3 & -2 & -1 & 0 & 1 & 2 \\ -2 & -1 & 0 & 1 & 2 & 3 \\ -1 & 0 & 1 & 2 & 3 & 4 \\ 0 & 1 & 2 & 3 & 4 & 5 \end{Bmatrix} \tag{4.25}$$

Ideally, the cross-response lattice derived from this array should have no lattice points with magnitude greater than one to insure minimum aliased signal power for every frequency shift and delay. Inspection of the particular $\{\phi_{ij}\}$ calculated above shows that the cross-response lattice corresponding to Γ_1 and Γ_2 does, in fact, not have any lattice point with a magnitude greater than one. However, the lattice points whose location is of primary concern are the ones for which $i = j$, since they represent aliased energy occuring simultaneously with the transmitted symbol. This aliased energy affects the probability of error in the same way as spreading of the symbol energy by a fluctuating channel. Assuming a "mark" is transmitted the aliasing which results from mistuning of the receiver adds to the thermal noise in the "space" receiver. In fact, assuming negligible thermal noise, the ratio of symbol energy to aliased energy will only be

$$\frac{E}{E_a} = \frac{PT}{PT/M} = M \ , \tag{4.26}$$

when

$$\phi = \phi_{ij} \ , \qquad j = 1, 2, \ldots, M \ .$$

Thus for short codes the self-noise caused by aliasing can dominate the thermal noise, even though the codes have been selected to give a low cross-ambiguity function. Consequently, the "mark" and "space" symbols should be chosen to place the lattice points on the $\tau = 0$ axis (which correspond to aliased energy present at the same time as the signal) at as large a ϕ as possible. Of way of constructing a pair of sequences with maximum distance to the points on the $\tau = 0$ axis (although not with the uniform cross-ambiguity property) is to choose an arbitrary sequence of length M (M even) and add M/2 to those elements of the sequence such that $\gamma_j \leq M/2$ or subtract M/2 from elements such that $\gamma_i > M/2$. This procedure results in a new sequence which is not only orthogonal but the first lattice points on the $\tau = 0$ axis occur for $|\phi| = M/2$. This procedure will maintain the portion of the $\tau = 0$ axis for which $|\phi| \ll M/2$ free of cross-ambiguity responses, and will minimize aliasing for small amounts of uncompensated frequency offset in the receiver.

As M increases the frequency offset Δf which can be considered small also increases since the condition which must be satisfied is

$$\phi = \frac{\Delta f}{T} M \ll \frac{M}{2} , \tag{4.27}$$

which is one more reason that M should be made as large as possible consistent with a low error rate.

In the next section the possibility of using more than two transmitter symbols will be considered.

4.7 M-ary Source Alphabets

In previous sections only binary source alphabets were considered. However, by increasing the number of orthogonal sequences in the source alphabet to M the number of binary bits encoded on each sequence will also increase to

$$\log_2 M \frac{\text{bits}}{\text{sequence}} . \tag{4.28}$$

Thus by using larger source alphabets more efficient use may be made of the channel. The penalty incurred by using larger source alphabets is a larger required sequence energy to maintain the same sequence error rate. Also, the mutual interference resulting from receiver mistuning will be increased since more sequences are sharing the same time-frequency space. However, since more bits are encoded on each sequence the energy per bit necessary to maintain a specified bit error rate will decrease. The larger number of source symbols also permits, instead of decreasing to energy per bit to hold the bit error rate constant, increasing the sequence duration by $\log_2 M$, which maintains the original bit rate and, for a peak power limited system, increases the received sequence energy by $\log_2 M$. This increase in sequence SNR will permit a corresponding increase in M which results in diversity improvement (as long as the diversity channel correlation remains small) and a reduction in effective fluctuation bandwidth. For a low data rate system these factors can provide a reduction of bit error rate in addition to that provided by the larger source alphabet.

In order to analyze the effect of M-ary encoding further, it will be useful to derive an upper bound on the encoded bit error probability in terms of the binary error probability derived previously. If the outputs of each receiver matched to a particular sequence Γ_i is given by B_i, then the probability that the receiver decides incorrectly that Γ_i, $i \neq M$, was sent when in fact Γ_M is the transmitted sequence is given by the probability of the union of events that $B_i > B_M$, $i \neq M$, i.e.,

$$P_{e,M} = \Pr\left\{\bigcup_{i=1}^{M-1} (B_i > B_M)\right\} . \tag{4.29}$$

By the union bound

$$P_{e,M} \leq \sum_{i=1}^{M-1} \Pr\{B_i > B_M\} \tag{4.30}$$

Since the symbol alphabet is symmetric

$$\Pr\left\{B_i > B_M\right\} \tag{4.31}$$

is just the binary probability of error P_e. Thus, the M-ary alphabet symbol error probability is bounded above by

$$P_{e,M} \leq (M - 1)\, P_e \ , \tag{4.32}$$

and for small sequence error probabilities this bound becomes tight. The sequence error probability can be related to the encoded bit error probability by recognizing that if a sequence is incorrectly received the average probability of error for the bits in the sequence is

$$\frac{1}{2}\left[\frac{2^{\log_2 M}}{2^{\log_2 M} - 1}\right] \tag{4.33}$$

Using this expression and Eq. (4.32) the probability that an encoded bit is received in error is bounded by

$$\frac{(M - 1)}{2}\left[\frac{2^{\log_2 M}}{2^{\log_2 M} - 1}\right] P_e \tag{4.34}$$

For large M the bit error probability is approximated by

$$P_{b,M} \leq \frac{M - 1}{2}\, P_e \ . \tag{4.35}$$

As an example of the benefit obtained by using large source alphabets, a sixteen symbol alphabet will be compared with the error probability obtained with binary encoding.

Assume that the received SNR is 13 dB and that $b = 0.64$. From Fig. 5.4(a) the probability of error for a binary alphabet using sequences of length 16 is 2×10^{-2}. If, instead, the source alphabet is increased to sixteen symbols, the sequence SNR can be increased to 19 dB without decreasing the transmission rate since $\log_2 16 = 4$, which allows the sequence duration to be increased by a factor of four. Using Fig. 5.4(c) with a diversity channel SNR of

$$19 - 10 \log_{10} 16 = 5 \text{ dB}$$

and the $b = 2.56$ curve (because the duration of each subpulse is also increased by a factor of four) the new sequence error rate is 7×10^{-5} and thus the bit error probability is bounded by

$$P_{b,M} \leq \frac{15}{2} \times 7 \times 10^{-5} = 5.25 \times 10^{-4} .$$

Comparing this bound with the error rate for the binary alphabet shows that the bit error rate has decreased from 2×10^{-2} to 5.25×10^{-4} with no reduction in information rate.

The next section will consider the construction of orthogonal alphabets.

4.8 Constructing Sets of Orthogonal Sequences

In discussing the advantages of using source alphabets larger than binary it was assumed that, in fact, these large source alphabets of orthogonal sequences could be constructed. In this section a simple algorithm for constructing sets of up to M sequences will be derived, and it will be shown that not only are the members of this set orthogonal but that the cross-response lattice points on the $\tau = 0$ axis are uniformly distributed.

Consider the set of sequences (for which $M = 6$)

$$\Gamma_1 = \{1,2,3,4,5,6\} ,$$

$$\Gamma_1 = \{1,2,3,4,5,6\} ,$$

$$\Gamma_2 = \{2,4,6,1,3,5\} ,$$

$$\Gamma_3 = \{3,6,2,5,1,4\} ,$$

$$\Gamma_4 = \{4,1,5,2,6,3\} ,$$

$$\Gamma_5 = \{5,3,1,6,4,2\} ,$$

$$\Gamma_6 = \{6,5,4,3,2,1\} .$$

Inspection of this set shows that Γ_i is obtained from Γ_1 by modulo M+1 multiplication by i and, therefore, every sequence can be derived by modulo M+1 multiplication from any other. A sequence length of six is only for illustrative purposes; M may be any integer provided that M+1 is prime. Were M+1 not prime then some of the Γ_i would have members equal to zero, which is not allowable.

It is easily verified that all of the Γ_i are orthogonal. Suppose that Γ_k is obtained from Γ_j by modulo M+1 multiplication by the integer L_k. Then, the elements of Γ_k are given by

$$\gamma_{i,k} = \left[\gamma_{i,j} \times L_k\right]_{M+1} \tag{4.36}$$

By the definition of modulo M+1 multiplication $\gamma_{i,k}$ can also be written

$$\gamma_{i,k} = (\gamma_{i,j} \times L_k) - Q_i(M+1) \tag{4.37}$$

where Q_i is an integer selected to satisfy the equality. But the sequences will be orthogonal if

$$\gamma_{i,k} - \gamma_{i,j} \neq 0 \ . \tag{4.38}$$

Subtracting $\gamma_{i,j}$ from the expression for $\gamma_{i,k}$ implies that

$$\gamma_{i,k} - \gamma_{i,j} = \gamma_{i,j} (L_k - 1) - Q_i(M+1) \ . \tag{4.39}$$

The only way for this expression to equal zero is if

$$\gamma_{i,j} = \frac{Q_i}{L_k - 1} (M+1) \ . \tag{4.40}$$

Since $\gamma_{i,j}$ must be an integer and since $M+1$ is prime the only values of $\gamma_{i,j}$ which solve this expression are integer multiples of $M+1$. But the $\gamma_{i,j}$ may only have integer values between 1 and M and, therefore, there is no allowable $\gamma_{i,j}$ which solves the expression. Thus all of the Γ_i are orthogonal.

Using a similar argument it can be shown that the distances

$$\gamma_{i,k} - \gamma_{i,j} = n_i \tag{4.41}$$

and

$$\gamma_{m,k} - \gamma_{m,j} = n_m \tag{4.42}$$

are unique. Recalling the definition of the cross-response lattice, this uniqueness means that there will be at most one lattice point located at each integral multiple of M/T. Since by the previous discussion n_i is given by

$$n_i = \gamma_{i,k} - \gamma_{i,j} = \gamma_{i,j} (k - 1) - Q_i(M+1) \tag{4.43}$$

then,

$$n_m = \gamma_{m,k} - \gamma_{m,j} = \gamma_{m,j} (L_k - 1) - Q_m(M+1) \ . \qquad (4.44)$$

Using the above expressions

$$n_i - n_m = (L_k - 1)(\gamma_{i,j} - \gamma_{m,j}) + (M+1)(Q_m - Q_i) \ . \qquad (4.45)$$

If

$$n_i - n_m = 0 \ ,$$

then

$$\gamma_{i,j} - \gamma_{m,j} = \frac{Q_i - Q_m}{L_k - 1} (M+1) \ . \qquad (4.46)$$

But, the above expression can only be satisfied if $\gamma_{i,j} - \gamma_{m,j}$ is an integer multiple of M+1. Since $\gamma_{i,j}$ and $\gamma_{m,j}$ are distinct members of the sequence Γ_j there will be no solution. Thus it follows that

$$n_i - n_m \neq 0 \ , \qquad (4.47)$$

which is the desired property.

Since every possible sequence of length M can be obtained from the sequence

$$\Gamma_M = \{1,2,3,\ldots,M\}$$

by interchanging the elements of Γ_M and since the conditions for orthogonality and low cross-response were independent of sequence order, it follows that any sequence will yield a set of M orthogonal sequences

with low cross-response under modulo $M + 1$ multiplication. For $M + 1$ not a prime number the maximum number of orthogonal sequences which may be constructed in this manner is equal to the number of positive integers less than or equal to M which are relatively prime to $M + 1$.

Since each sequence may be used to construct a set of M (for $M+1$ prime) orthogonal sequences, there are $(M - 1)!$ distinct orthogonal sets which can be constructed by this method.

There are other ways in which sets of orthogonal sequences may be generated, but in general these sets will not have a desirable cross-response. As an example of one of these methods, let

$$\Gamma_1 = \{1,2,\ldots,M\}$$

and add any positive integer $\leq M - 1$ to the elements where, if the addition results in an element $\geq M + 1$, the value of that element will be defined to be its value taken modulo M. Although M orthogonal sequences result, the cross-response for

$$\Gamma_2 = \{2,3,\ldots,M,1\}$$

has a value of $M - 1$ located at $\tau = 0$, $\phi = 1/T$, which would result in a large sensitivity to mistuning.

This discussion concludes Chapter IV. The emphasis in Chapter V will be on implementation of the CFSK receiver and acquisition and coding techniques.

Chapter V

RECEIVER IMPLEMENTATION, CODING, AND FREQUENCY ACQUISITION

5.1 Implementation of the CFSK Receiver

The general form of the CFSK receiver is shown in Fig. 5-1. The filters in each branch are matched to a particular subpulse of the coded frequency sequence. The filters are followed by square-law detectors which perform the incoherent detection. After each of the square-law detectors is a series of delays which serve to delay each subpulse to bring it into correct time alignment with the other subpulses of the same sequence. If the sequences are orthogonal, when Γ_j is transmitted only the delayed filter output sum B_j will have the subpulses aligned. The remaining B_i will have the subpulses dispersed in time according to the cross-ambiguity function of B_j and B_i. Then the summed delayed filter outputs are compared at time $t = T_0$ and the largest assumed to be the transmitted sequence.

Functionally, the receiver has five types of components: subpulse filters, square-law detectors, delay lines, summing elements, and decision logic which decides which sequence was most likely to have been transmitted. All of these functions may be implemented either with analog elements or digitally; however, in the case of the delay lines a digital implementation appears to be the only feasible alternative.

One possible hybrid receiver uses analog filters and square-law detectors followed by an A-D converter which permits the remainder of the receiver to be implemented digitally. For sequences with total duration of one second and M on the order of ten to one-hundred analog filters with center frequencies of up to 10 kHz and 100 Hz bandwidth would be required. Although stable analog filters with the required bandwidth can be constructed, such a receiver would not allow changes in the number of frequencies and sequence duration without considerable modification.

An entirely digital receiver has several advantages over the hybrid receiver: completely stable, high-Q filters may be synthesized, the

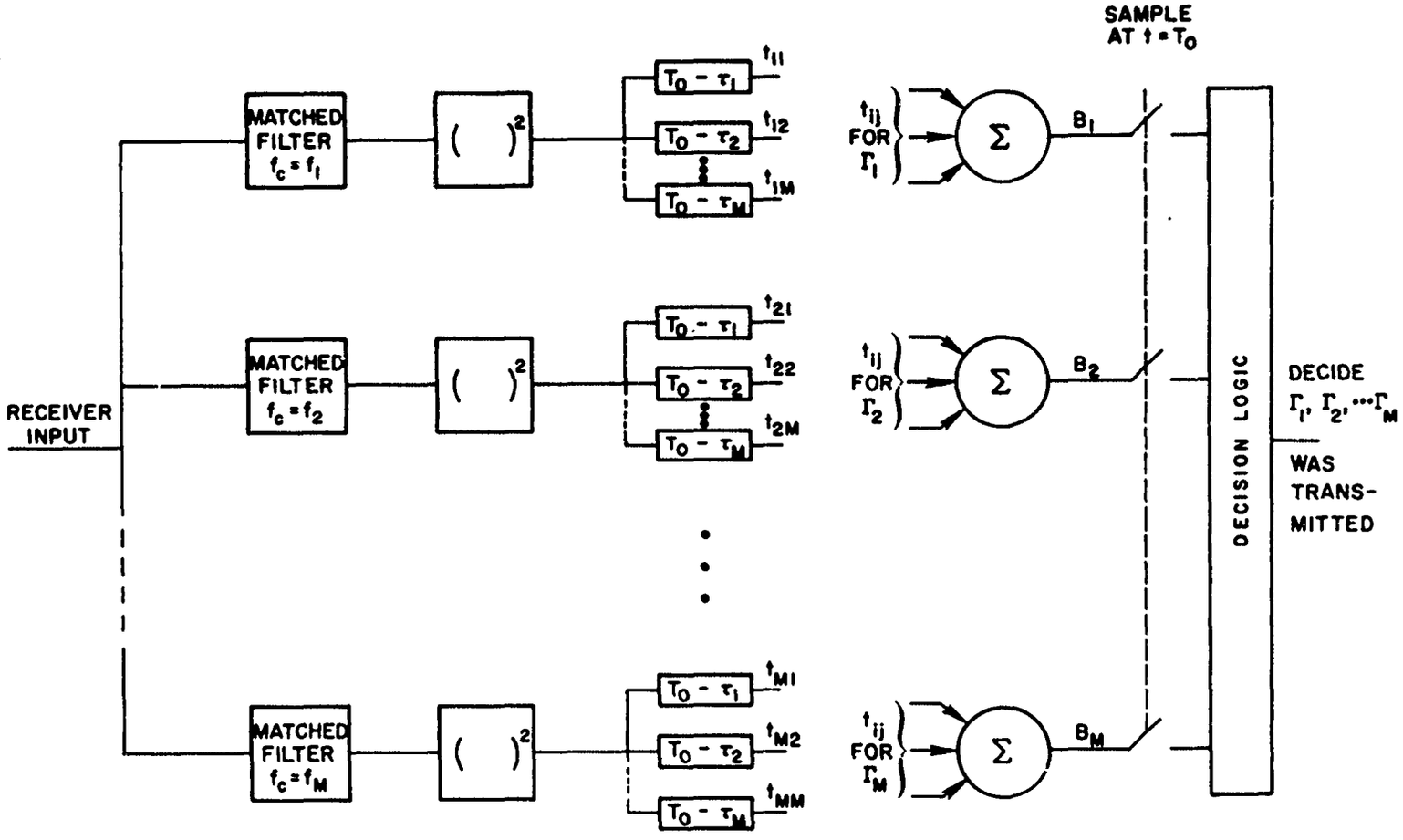


Fig. 5.1. BLOCK DIAGRAM OF CFSK SEQUENCE RECEIVER.

nonlinear detectors are easily implemented and stable, accurate time delays on the order of seconds are realizable. In addition, only software changes are necessary to permit detection of sequences with different coding, number of frequencies, and duration.

To enable as low a sample rate as possible the received data would be translated to a near lowpass spectrum. This low frequency data would be sampled and quantized in an A-D converter and stored. Segments of this data one subpulse in duration are Fourier transformed to obtain the power spectrum at each of the possible subpulse frequencies. The power at particular frequencies and delays is then added to align the subpulses of each of the possible sequences in time. These sums are then compared and the largest assumed to correspond to the transmitted sequence. Since the total bandwidth will usually be less than several kHz, it will be possible to perform these computations in real time.

Because of the inherent flexibility of the digital implementation it appears that a completely digital receiver rather than a hybrid is preferred.

5.2 Frequency Acquisition of CFSK Sequences

Since the sequence bandwidth can be a small fraction of the range of frequency uncertainty resulting from long-term oscillator drift and unknown doppler shift, before data can be received the receiver must be correctly tuned.

The simplest procedure is to tune the receiver in steps of approximately $M/2T$ across the range of frequency uncertainty. Since the duration of each sequence is T a range of frequency uncertainty ΔF may be covered in

$$\Delta T = (2T/M)(T\Delta f) \tag{5.1}$$

with a high probability of detection.

A narrowband transmission of duration $1/T$ and equal energy would require approximately M times longer for acquisition since the bandwidth is a factor of M less than the CFSK sequences.

For certain sequences acquisition can be even more rapid. Recalling the incoherent channel response function for a linear frequency progression, it is apparent that for $|\emptyset T| \leq M/2$ the SNR loss from mistuning will be less than 3 dB, which permits the tuning to be incremented in steps of $M^2/4T$ instead of $M/2T$. Since the total SNR is assumed adequate to obtain a $P_{\epsilon,M}$ of 10^{-3} or better, even with a 3 dB loss from misturning the detection and false alarm probabilities will be satisfactory. With these assumption the acquisition time can be reduced to

$$\Delta T = 4T^2 \Delta f / M^2 \quad . \tag{5.2}$$

Thus for large M and linear frequency sequences the acquisition time can be reduced by a factor of $M^2/2$ from the time required by narrowband transmissions.

This ability to detect the presence of signal energy even though the receiver is severely mistuned could considerably reduce the amount of computation required for the detection of weak echoes such as encountered in radar astronomy because of the reduction in the number of doppler channels which must be searched.

The next section will consider the advantages obtained with simple error correcting codes.

5.3 Coding for Reliable Transmission in M-ary Systems

For certain purposes the error rates obtained with CFSK sequences will not be sufficiently small and error detecting and possibly error correcting will be required. Because incorrect reception of an M-ary symbol results in approximately one-half of the bits associated with that particular symbol being in error, standard serial encoding would require very long code blocks to be sent for efficient transmission, particularly for large M.

An alternative is longitudinal coding where the binary symbols to be encoded as an M-ary symbol (where $\log_2 M$ is an integer) is to arrange the binary data to be transmitted in an array, with the last p

rows being error detecting and correcting codes, and whose rows, except for the last p are the binary data to be transmitted.

Then the rows are sequentially encoded as one of the M sequences. The error detecting and correcting codes are also encoded and transmitted. At the receiver the reverse column by column decoding is performed and any needed corrections performed.

Suppose that P data sequences followed by p error correcting sequences are sent. Then, the information rate drops to a fraction

$$\frac{P}{p + P} \tag{5.3}$$

or its original value.

However, errors will cluster in a particular M bit word as a result of a sequence error. Since it requires 2 bits to correct a single error, an incorrect sequence can be corrected with 2 redundant sequences. Thus to make an uncorrectable error 2 sequences out of the $P + p$ sequences must be in error.

Assuming the sequence errors to be independent with probability $P_{e,M}$ gives

$$\Pr\{N \text{ errors in } P + p \text{ sequences}\} = \binom{P+p}{N} P_{e,M}^{N} \left(1 - P_{e,M}\right) p + P - N \tag{5.4}$$

The probability of 2 errors is then

$$\frac{1}{2} (P + p)(p + P - 1)\, P_{e,M}^{2}(1 - P_{e,M})\, p + P - 2 \tag{5.5}$$

which, if $p + P \gg 1$ and $P_{e,M} \ll 1$, becomes

$$\Pr\{2 \text{ errors in } P + p \text{ sequences}\} \approx \frac{1}{2} (p + P)^{2} P_{e,M}^{2} . \tag{5.6}$$

This expression shows that if

$$P \approx \frac{1}{P_{e,M}} , \tag{5.7}$$

then the probability of an uncorrectable error is large and the error control codes are useless. As long as P is sufficiently small, however, error control coding can substantially reduce the error rate as the following example will show.

Suppose that $P_{e,M} = 10^{-4}$ and $p + P = 12$, then two sequences out of the $p + P$ will be detected incorrectly with a probability of 6.6×10^{-7}. Since there are $\log_2 M$ bits encoded on each sequence and two sequence errors imply that on the average $\log_2 M$ bit errors will result out of the $(p + P) \log_2 M$ bits in the code block, the bit error rate becomes (assuming more than two sequence errors to have low probability)

$$P_b \approx \frac{\Pr\{\text{ 2 sequences in error}\} \log_2 M}{(p + P) \log_2 M} \tag{5.8}$$

$$\approx \frac{P}{2} P_{e,M}^2 , \qquad p \ll P . \tag{5.9}$$

This bit error rate is independent of M for $M \gg 1$.

For the values used in the previous example the bit error rate is approximately 5×10^{-8}. The bit error probability without error control is

$$P_b \approx \frac{M}{2} P_{e,M} , \tag{5.10}$$

which would result, for $M = 16$, in $P_b = 1.6 \times 10^{-3}$. Thus even relatively simple error control can give quite considerable improvement in error rates.

The disadvantage of this type of error control is that a substantial amount of storage is required both at the transmitter and receiver. However, since reliable, large-scale memories with low power consumption using MOS techniques are available, providing the required storage should not be a major problem.

Chapter VI

RECOMMENDATIONS FOR FURTHER INVESTIGATION

In the previous chapters, we have discussed the analysis and applications of coded frequency shift keyed (CFSK) sequences, with emphasis on their use in fluctuating channels. However, in the analysis we assumed that the frequency step spacing was equal to the inverse of the subpulse duration. A generalization to permit representation by the response lattice of sequences with non-harmonic spacing and possibly non-contiguous subpulses would have wide applications.

Another subject for further consideration would be the conditions for synthesizing long sequences with certain properties (for example, the "thumbtack" property) by combining shorter sequences.

In calculating the bounds on the probability of error for CFSK sequences, the effects of subpulse correlation were neglected as was the dependence of the sequence error probability on the actual form of the sequence. A more exact expression would be valuable in assessing the relative performance of different sequences.

Finally, experimental verification for these results either in the laboratory or under field conditions which permits realistic assessment should be performed. If these experiments verify the theoretical results obtained in this investigation, a more severe test such as use in a solar radar would confirm their utility.

BIBLIOGRAPHY

Barna, A., Pseudorandom Frequency Modulation in Range--Doppler Radar, Rep. SEL-68-046 (Sci. Rep. No. 29, NASA Grant NsG-377), Stanford Electronics Laboratories, Stanford, Calif., May 1968.

Bello, P. A., and Nelin, B. D., Predetection Diversity Combining with Selectively Fading Channels, IRE Trans. Commun. Systems, vol. CS-10, pp.32-42, March, 1962.

Bello, P. A., and Nelin, B. D., The Influence of Fading Spectrum on the Binary Error Probabilities of Incoherent and Differentially Coherent Matched Filter Receivers, IRE Trans. Commun. Systems, vol. CS-10, pp. 160-168, June, 1962.

Bello, P. A., and Nelin, B. D., The Effect of Frequency Selective Fading on the Binary Error Probabilities of Incoherent and Differentially Coherent Matched Filter Receivers, IEEE Trans. Commun. Systems, vol. CS-11, pp. 170-186, June, 1963; corrections in IEEE Trans. Comm. Tech., vol. COM-12, pp. 230-231, December, 1964.

Cook, C. D, and Bernfeld, M., "Radar Signals," Academic Press, New York, 1967.

Costas, J. P., Synchronous Communication, Proc. IRE, vol. 44, pp. 1713-1718. December, 1956.

Davenport, W. B. Jr., and Root, W. L., "Introduction to Random Signals and Noise," McGraw-Hill Book Company, New York, 1958.

Develet, J. A., A Threshold Criterion for Phase-Lock Demodulation, Proc. IRE, vol. 51, pp. 349-356, February 1963.

Evans, J. V., and Hagfors, T., "Radar Astronomy," McGraw-Hill Book Co., New York, N. Y., 1968.

Ferguson, M. J., Communication at Low Data Rates: Spectral Analysis Receivers, Philco-Ford Corp., Palo Alto, Calif., Tech. Mem. 124, Sept. 5, 1967.

Goldstein, R. M., et al, The Superior Conjunction of Mariner IV, Jet Propulsion Laboratory, Tech. Rept. 32-1092, April 1, 1967.

Golomb, S. W. (ed.), "Digital Communications with Space Application," Prentice-Hall, Inc., Englewood Cliffs, N. J., 1964.

Jordan, D. B., Greenberg, H., Eldredge, E. E., and Serniuk, W., Multiple Frequency Shift Teletype Systems, Proc. IRE, vol. 43, pp. 1647-1655, November, 1955.

Kuhn, B. G., et al., Orthomatch Data Transmission System, IEEE Trans. Space Elect. and Telem., vol. SET-9, September, 1963.

Marcum, J. I., A Statistical Theory of Target Detection by Pulsed Radar, IRE Trans. Inform. Theory, vol. IT-6, pp. 59-267, April, 1960; earlier Rand Corp. Memos RM-753, July, 1948 and RM-754, December, 1947.

Nuttall, A. H., Error Probabilities for Equicorrelated M-ary Signals under Phase-Coherent and Phase-Incoherent Reception, IRE Trans. Inform. Theory, vol. IT-8, pp. 305-314, July, 1962.

Pierce, J. N., Theoretical Diversity Improvement in Frequency Shift Keying, Proc. IRE, vol. 46, pp. 903-910. May, 1958.

Pierce, J. N., Multiple Diversity with Non-Independent Fading, Proc. IRE, vol. 49, pp. 363-364, January, 1961.

Pierce, J. N., Theoretical Limitations on Frequency and Time Diversity for fading Binary Transmissions, IEEE Trans. Commun. Systems, vol. CS-11, pp. 186-187, June, 1963; earlier, Air Force Cambridge Research Laboratory Report ERD-TR-60-169, July, 1960.

Pierce, J. N., Approximate Error Probabilities for Optimal Diversity Combining, IEEE Trans. Commun. Systems, vol. CS-11, pp. 352-354, September, 1963.

Pierce, J. N., and Stein, S., Multiple Diversity with Non-Independent Fading, Proc. IRE, vol. 48, pp. 89-104, January, 1960.

Price, R., and Green, P. E. Jr., A Communication Technique for Multipath Channels, Proc. IRE, Vol. 46, pp. 555-570, March, 1958.

Price, R., and Green, P. E. Jr., Signal Processing in Radar Astronomy--Communication via Fluctuating Multipath Media, M.I.T. Lincoln Lab., Tech. Rept. 234, October, 1960.

Price, R. and Hofstetter, E. M., Bounds on the volume and height distributions of the ambiguity function, IEEE Trans. IT-11, 207-214 (1965).

Proakis, J. G., Drouilhet, P. R., and Price, R., Performance of Coherent Detection Systems Using Decision-Directed Channel Measurement, IEEE Trans. Commun. Systems, vol. CS-12, pp. 54-63, March, 1964.

Rubin, W. L., and DiFranco, J. V., The effects of Doppler dispersion on matched-filter performance, Proc. IRE (Correspondence) 50, 2127-2128 (1962).

Schmidt, A. R., A Frequency Stepping Scheme for Overcoming the Disastrous Effects of Multipath Distortion on High-Frequency FSK Communications Circuits, IRE Trans. Commun. Systems, Vol. CS-8, pp. 44-47, March, 1960.

Shapiro, I. I., Effects of General Relativity on Interplanetary Time-Delay Measurements, M.I.T. Lincoln Lab., Tech. Rept. 368, DDC No. 614232, Lexington, Mass., 1964.

Sussman, S. M., A Matched Filter Communication System for Multipath Channels, IRE Trans. Inform. Theory, vol. IT-6, pp. 367-373, June, 1960.

Tausworthe, R. C., Theory and Practical Design of Phase-Locked Receivers, Vol. I, Jet Propulsion Laboratory, Tech. Rept. 32-819, February 15, 1966.

Turin. G. L., Communication through Noisy, Random-Multipath Channels, IRE Nat. Conv. Record, part 4, pp. 154-166, March, 1956.

Turin, G. L., On Optimal Diversity Reception, IRE Trans. Inform. Theory, vol. IT-7, pp. 154-166, July, 1961.

Turin, G. L., On Optimal Diversity Reception, II, IRE Trans. Commun. Systems, vol. CS-10, pp. 22-31, March, 1962.

Woodward, P. M., "Probability and Information Theory, with Applications to Radar," Pergamon Press, Oxford, 1953.

Viterbi, A. J., "Principles of Coherent Communication," McGraw-Hill Book Co., New York, N.Y., 1966.

REFERENCES

P. A. Bello and B. D. Nelin, The Influence of Fading Spectrum on the Binary Error Probabilities of Incoherent and Differentially Coherent Matched Filter Receivers, IRE Trans. Commun. Systems, vol. CS-10, pp. 160-168, June, 1962.

C. E. Cook, and M. Bernfeld, "Radar Signals," Academic Press, New York, 1967.

J. V. Evans and T. Hagfors, "Radar Astronomy," McGraw-Hill Book Co., New York, N. Y., 1968.

R. M. Goldstein, et al, The Superior Conjunction of Mariner IV, Jet Propulsion Laboratory, Tech. Rept. 32-1092, April 1, 1967.

J. N. Pierce, Theoretical Diversity Improvement in Frequency-Shift Keying, Proc. IRE, vol. 46, pp. 903-910, May, 1958.

R. Price, and P. E. Green, Jr., A Communication Technique for Multipath Channels, Proc. IRE, Vol. 46, pp. 555-570, March, 1958.

R. Price, and P. E. Green, Jr., Signal Processing in Radar Astronomy--Communication via Fluctuating Multipath Media, M.I.T. Lincoln Lab., Tech. Rept. 234, October, 1960.

I. I. Shapiro, Effects of General Relativity on Interplanetary Time-Delay Measurements, M.I.T. Lincoln Lab., Tech. Rept. 368, DDC No. 614232, Lexington, Mass., 1964.

R. C. Tausworthe, Theory and Practical Design of Phase-Locked Receivers, Vol. I, Jet Propulsion Laboratory, Tech. Rept. 32-819, February 15, 1966.

G. L. Turin, On Optimal Diversity Reception, II, IRE Trans. Commun. Systems, vol. CS-10, pp. 22-31, March, 1962.

A. J. Viterbi, "Principles of Coherent Communication," McGraw-Hill Book Co., New York, N. Y., 1966.

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R & D

Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author)
Stanford Electronics Laboratories
Stanford University
Stanford, California

2a. REPORT SECURITY CLASSIFICATION
UNCLASSIFIED

2b. GROUP

3 REPORT TITLE
CODED FREQUENCY SHIFT KEYED SEQUENCES WITH APPLICATIONS TO LOW DATA RATE COMMUNICATION AND RADAR

4 DESCRIPTIVE NOTES (Type of report and inclusive dates)
Technical Report and Ph.D. Dissertation

5 AUTHOR(S) (First name, middle initial, last name)
Michael Jon Sites

6 REPORT DATE
September 1969

7a. TOTAL NO. OF PAGES

7b. NO OF REFS

8a CONTRACT OR GRANT NO
Nonr-225(83), NR 373 360

b PROJECT NO

c

d

9a. ORIGINATOR'S REPORT NUMBER(S)
SU-SEL-69-033
TR No. 3606-5

9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)

10 DISTRIBUTION STATEMENT
This document has been approved for public release and sale; its distribution is unlimited.

11 SUPPLEMENTARY NOTES

12. SPONSORING MILITARY ACTIVITY
U.S. Navy
Office of Naval Research
Washington, D.C.

13 ABSTRACT

Coded discrete frequency sequences can provide greatly improved performance over conventional techniques when the fluctuation bandwidth of the communication channel is a significant fraction of the transmission bandwidth. These fluctuations result from medium, equipment, and, in the case of radar, random target variations.

The ability to reliably detect these sequences under frequency shift is investigated using a simple algorithm to calculate an approximation of the true ambiguity function. This investigation leads to certain necessary conditions and a set of coupled equations which permit a sequence to be synthesized from the ambiguity function approximation.

Error rates for CFSK sequences are calculated and criteria for optimizing the performance in rapidly fluctuating channels given. An algorithm for constructing sets of orthogonal sequences with desirable cross-ambiguity properties is developed and the performance of these sets of orthogonal sequences compared with binary code alphabets.

Finally, consideration is given to digital and analog implementations of the special receiver required for the CFSK sequences, error correction coding problems associated with M-ary encoding, and acquisition behavior.

DD FORM 1 NOV 65 1473 (PAGE 1)

S/N 0101-807-6801

UNCLASSIFIED

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT